

human	Motif 0	
tez1	AKFLHMLMSVYVVELLRSEFFVYVTEFTTFQKNR	
EST2	ISEIEWLVGKRSNAKMCLSDFEKRQIFAEFIWLYNSFIPIILQSSFFYITESSDLRNR	
p123	LKDFRWLFISD---IWFTKHNFNENLNQLAICFISWLFRLPIKIIQTFFYCYTEISSTVT- TREISWMQVET-SAKHFYYFDHEN-IYVLWKLRLWFEDLVVSLRRCFFYVTEQQKSYSK * *** **	
human	Motif 1	
tez1	LFYRKSVMSKLSIGIRQHLLKRVQLRDVSEAEVRQHREARPALLTSLRFRIPKP--DGL	
EST2	TVYFRKDIWKLRCRPF--TSMQMEAFKINENNNVRMDIQK-TTLPPAVIRLLPKK--NTF	
p123	IVYFRHDTWNLITPFIIVEYFKTYLVENNVCRNHSYTLS--NFNHSKMRILPKKSNNEF TYYYRKNIWDVIMKMSI-ADLKKEITLAEVQKEVEEWKKS-LGFAPGKLRLLIPKK--TTF * * *	
human	Motif 2	
tez1	RPIVNMDYVVGARTFRREKRAERLTSRVKALF-SVLNVERA	
EST2	RLITN-LRKRFLIKMGSNKKMLVSTNQTLRPVASILKHLINEESSGIPFNLEVVMKLLTF	
p123	RIIAIPCRGADEEEFTIYKENHKNAIQPTQKILEYLRNKRPTSFTKIYSPQTADRIKEF RPIMTFNKKIVNSDRKTTTTLTNTKLLNSHLMLKTLKN-RMFKDPPFGFAVFNYYDDVMKKY * * *	
tez1	Motif 3 (A)	
EST2	KKDLLKHRMFGR-KKYFVRIDIKSCYDRIKQDLMFRIYVKK-KLKDPEFVIRKYATIHATS	
p123	KQRLKKFNINVLPELYFMKFDVKSCYDSIPMECMRIILKD-ALKNENGFFVRSQYFFNTN EEFVCKWKQVGQPKLFFATMDIEKCYDSVNREKLSSTFLKTTKLLSSDFWIMTAQILKRKN * * *	

FIG. 1

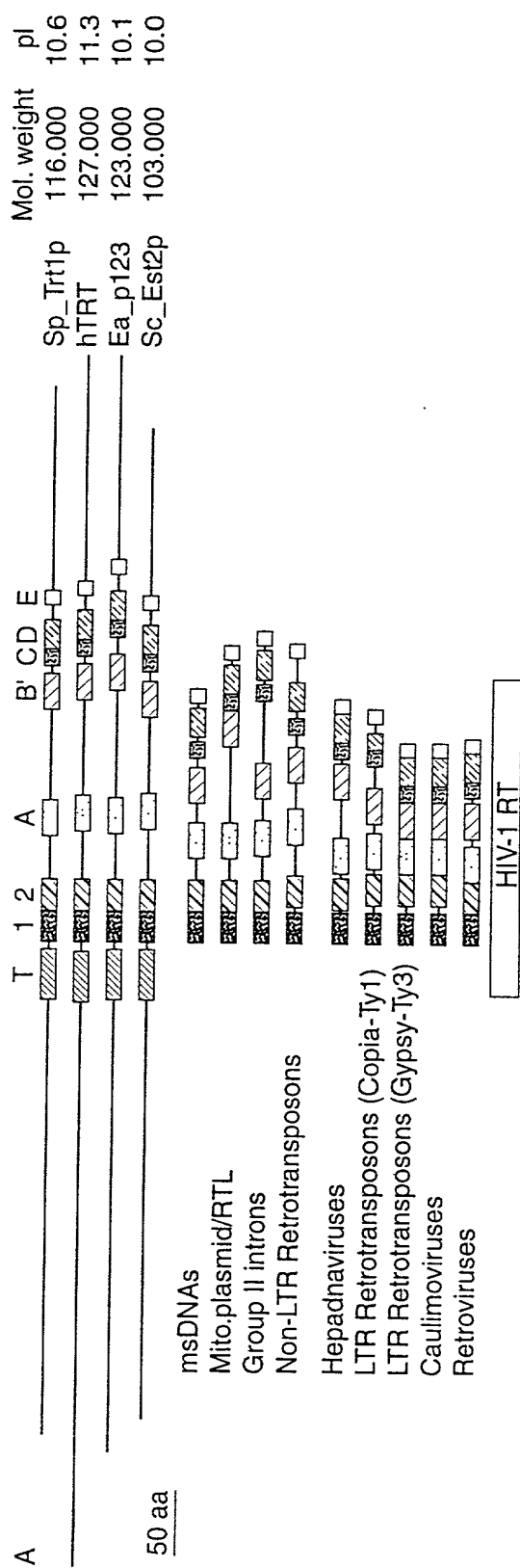


FIG. 2

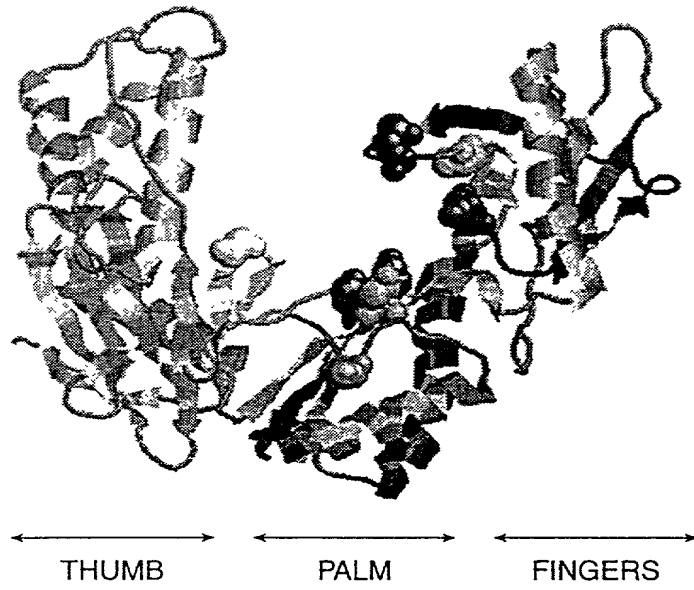


FIG. 3

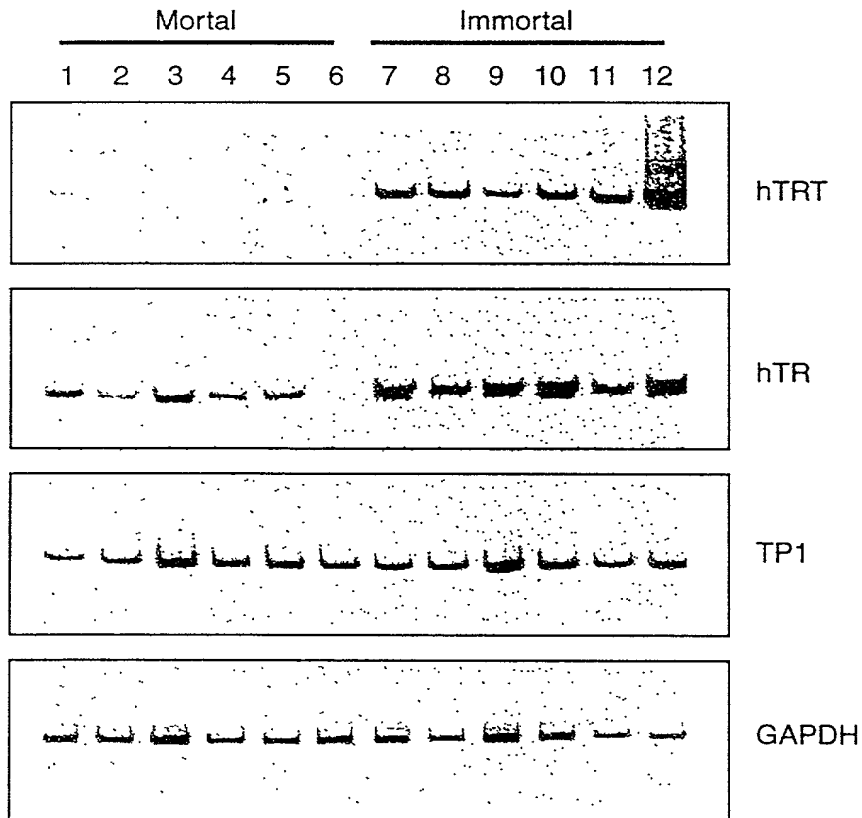


FIG. 5

Motif T		Motif 1		Motif 2		Motif A	
TRT con	WL hh hh pFFY TE p p Y RK W L h I K	h hrhPKK p	FRhi h h K	0 FRLTNLRKRFLIKMGSNKMLVSTNQTL	40 PCLYFh hDh CYD I hhK K	82	
Sp_Trt1p	429 WLYNSFIIPILQFFFYITESSDLNRRTVFRKDIWKLRCRPFITSMKM 8	h hrhPKK p	FRhi h h K	0 FRLTNLRKRFLIKMGSNKMLVSTNQTL	40 PCLYFh hDh CYD I hhK K	82	
hTRT	546 WLMSVYVVELLRFFFYVTEFTQKRLFFRKSVWSKLQSIGIRQHLK 10	h hrhPKK p	FRhi h h K	0 FRLTNLRKRFLIKMGSNKMLVSTNQTL	40 PCLYFh hDh CYD I hhK K	82	
Ea_p123	441 WIFEDLVSLIRCFYFVTEQQSKYSKYFYRKNINWDIMKMSIADLKK 8	h hrhPKK p	FRhi h h K	0 FRLTNLRKRFLIKMGSNKMLVSTNQTL	40 PCLYFh hDh CYD I hhK K	82	
Sc_Est2p	366 WLFRQLIPKIIQTFYCYCTEISSTVT-IVFRHDTWNKLITPFIVEYFK 8	h hrhPKK p	FRhi h h K	0 FRLTNLRKRFLIKMGSNKMLVSTNQTL	40 PCLYFh hDh CYD I hhK K	82	
Motif B'		Motif C		Motif D		Motif E	
TRT con	K Y Q GIPQGS LS hL h Y DL	LLRL DDFLhIT	A F h G c p N cK	0 AKKFLNLSLRGFEKHNFSSTSEKTVI	17 KKRMPFFGFSV	181	
Sp_Trt1p	429 WLYNSFIIPILQFFFYITESSDLNRRTVFRKDIWKLRCRPFITSMKM 8	LLRL DDFLhIT	A F h G c p N cK	0 AKKFLNLSLRGFEKHNFSSTSEKTVI	17 KKRMPFFGFSV	181	
hTRT	546 WLMSVYVVELLRFFFYVTEFTQKRLFFRKSVWSKLQSIGIRQHLK 10	LLRL DDFLhIT	A F h G c p N cK	0 AKKFLNLSLRGFEKHNFSSTSEKTVI	17 KKRMPFFGFSV	181	
Ea_p123	441 WIFEDLVSLIRCFYFVTEQQSKYSKYFYRKNINWDIMKMSIADLKK 8	LLRL DDFLhIT	A F h G c p N cK	0 AKKFLNLSLRGFEKHNFSSTSEKTVI	17 KKRMPFFGFSV	181	
Sc_Est2p	366 WLFRQLIPKIIQTFYCYCTEISSTVT-IVFRHDTWNKLITPFIVEYFK 8	LLRL DDFLhIT	A F h G c p N cK	0 AKKFLNLSLRGFEKHNFSSTSEKTVI	17 KKRMPFFGFSV	181	
Motif B'		Motif C		Motif D		Motif E	
RT con	hPQG pP hh h	h Y Ddhhh	Gh h cK h	0 AKKFLNLSLRGFEKHNFSSTSEKTVI	17 KKRMPFFGFSV	181	
Sc_a1	TYHKPMLGLPQGSLSIPILCNIVMTLVDNWLEDYI	h Y Ddhhh	Gh h cK h	0 AKKFLNLSLRGFEKHNFSSTSEKTVI	17 KKRMPFFGFSV	181	
Dm_TART	RAGQIGAGVPQGSNLGPILYSIFSSDMPLPHIYHP	h Y Ddhhh	Gh h cK h	0 AKKFLNLSLRGFEKHNFSSTSEKTVI	17 KKRMPFFGFSV	181	
HIV-1	GIRYQYNVLPQGSNKGSPAIFQSSMTKILEPKKQN	h Y Ddhhh	Gh h cK h	0 AKKFLNLSLRGFEKHNFSSTSEKTVI	17 KKRMPFFGFSV	181	

FIG. 4

5/34

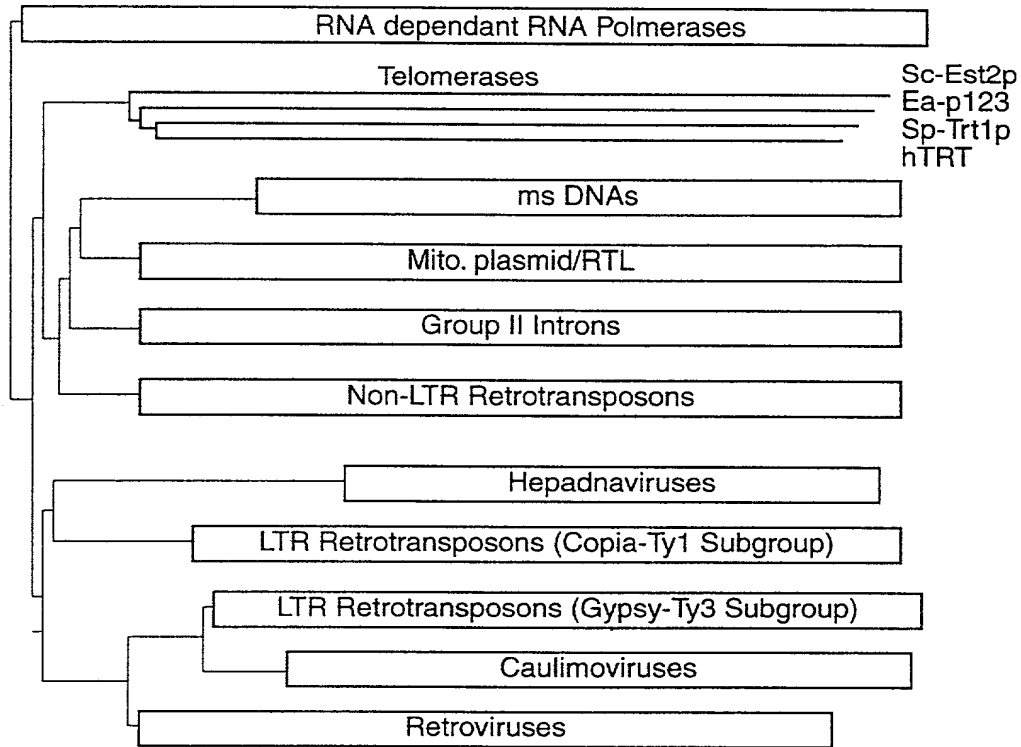


FIG. 6

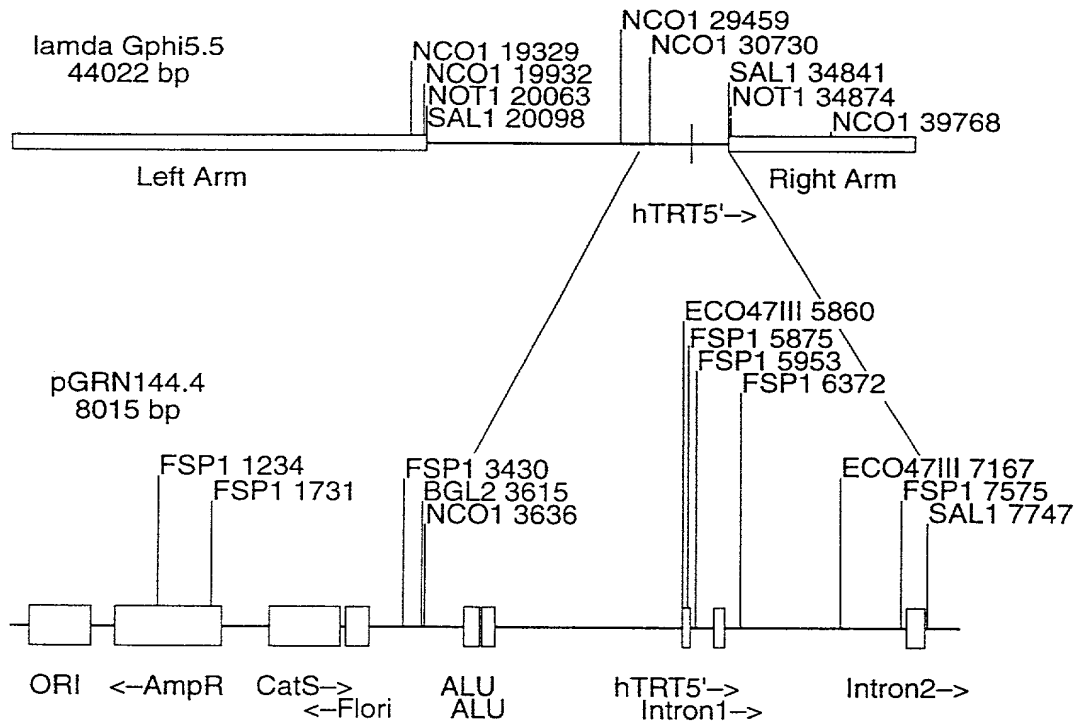


FIG. 7

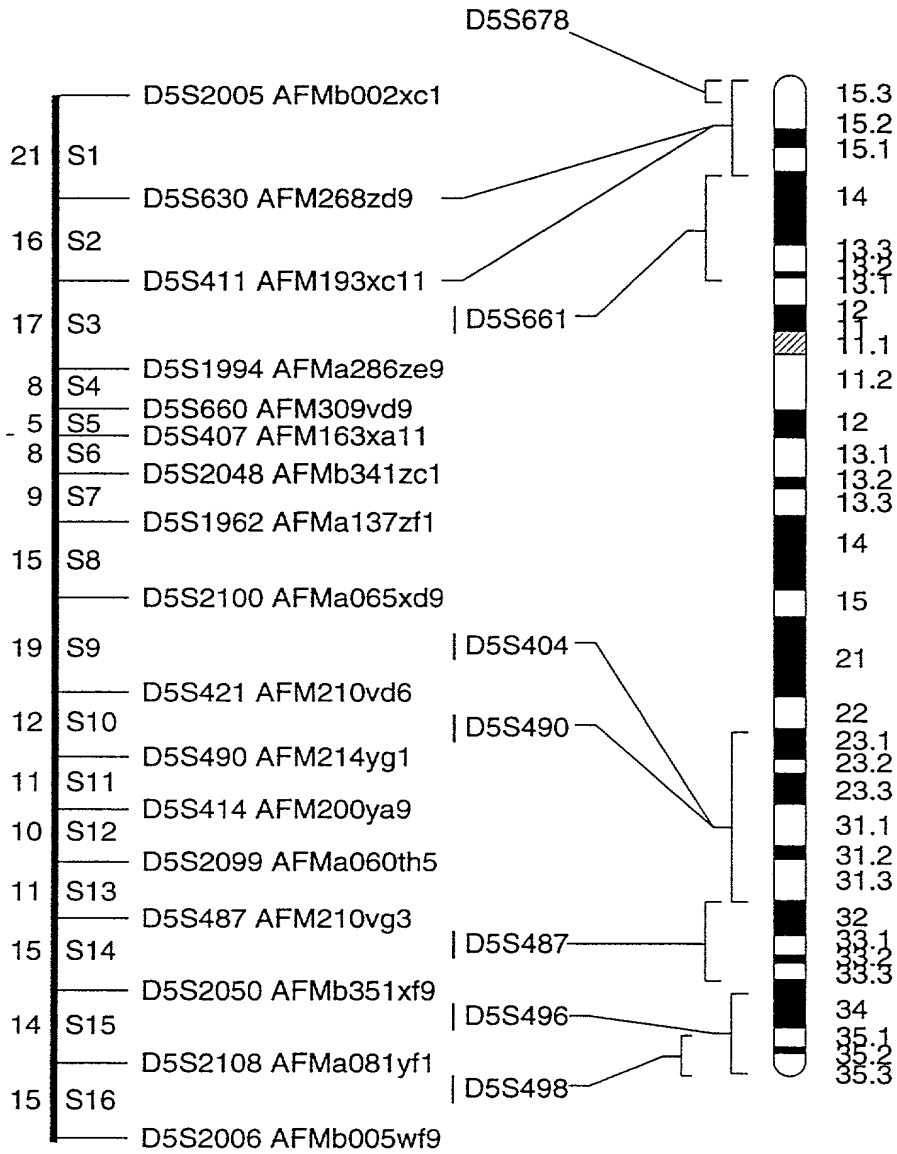


FIG. 8

7/34

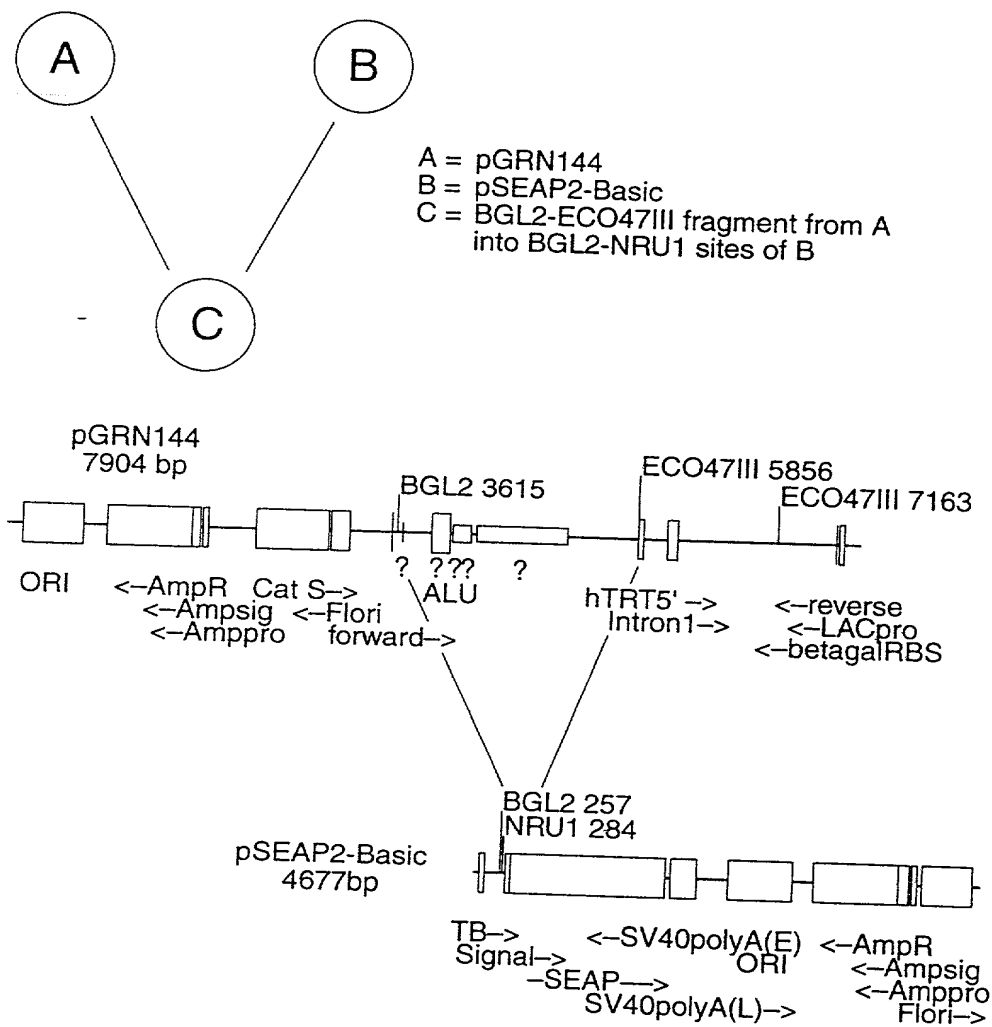


FIG. 9

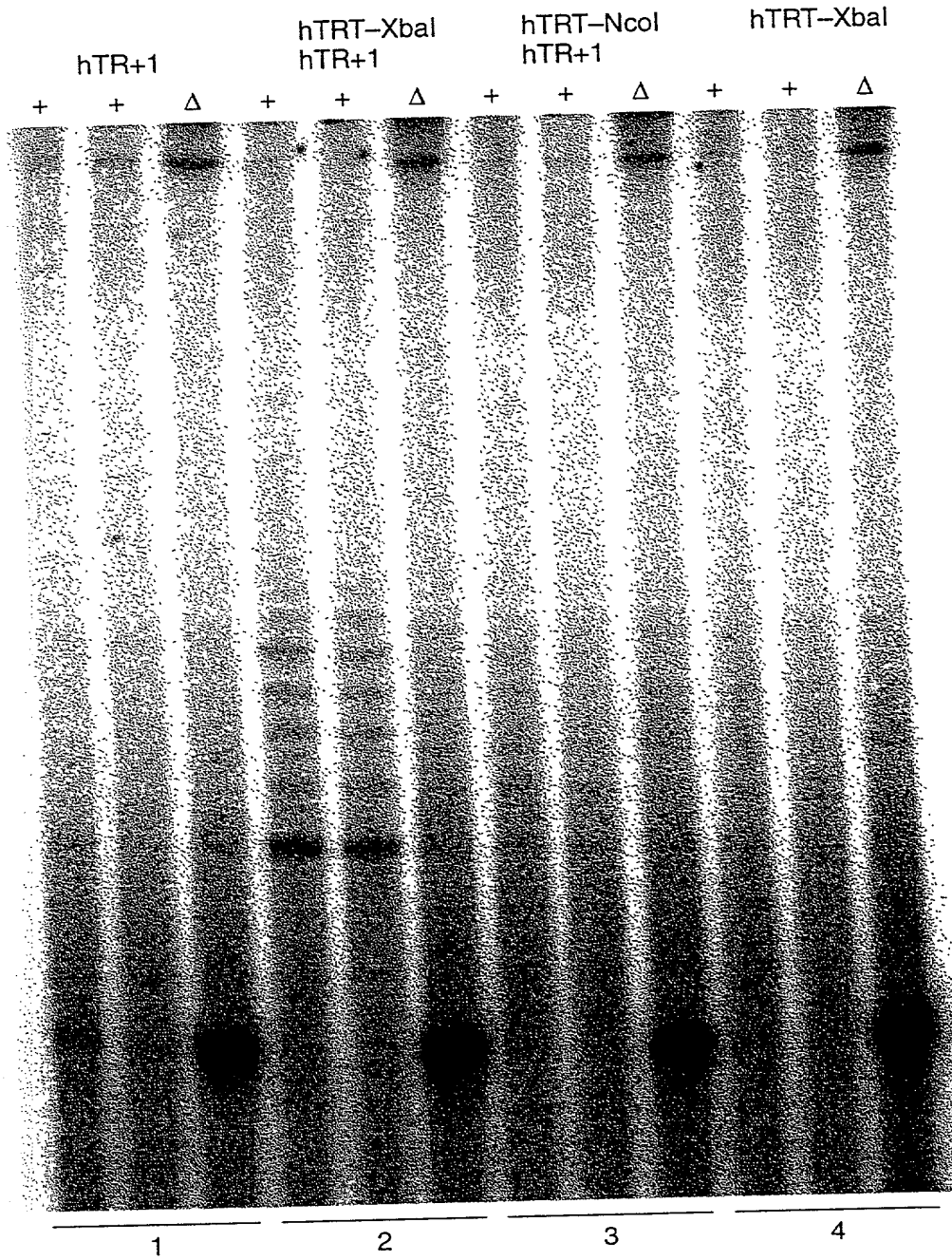


FIG. 10A

9/34

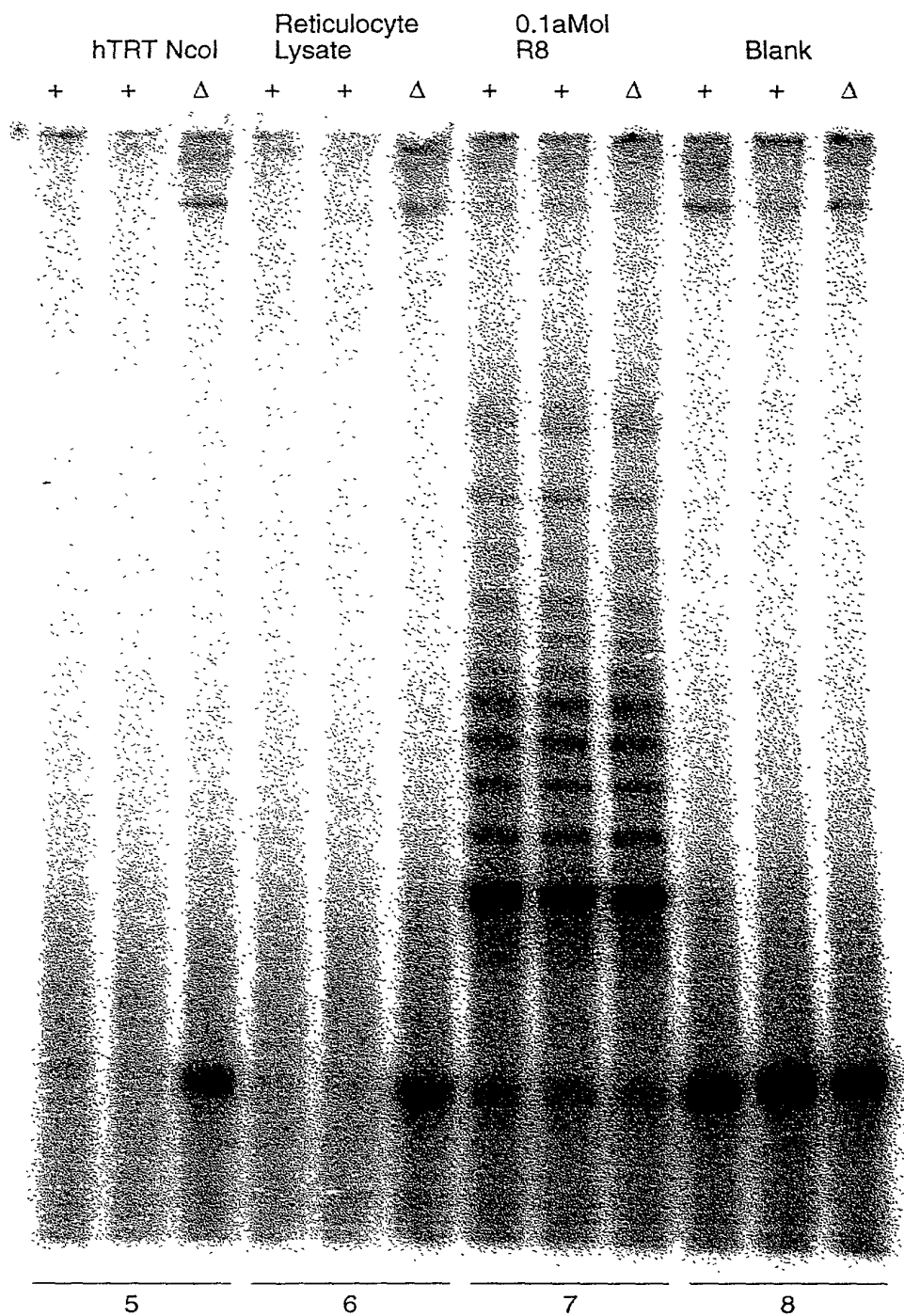


FIG. 10B

Telomerase Specific Motifs

	MOTIF T	MOTIF T'
TRT con	W1	
hTRT	546 WLMSSVVELLRSSFFVVTETTFQKNRLFFYRKSWSKLQSIGI	13 EAEVR
spTRT	429 WLYNSFIIPILQSFYITESSDLNRRTVYFRKDIWKLLCRPFI	12 ENNV
Ea_p123	441 WIFEDLVSLIRCFYVTEQQKSYKTYRKNWDVIMKMSI	12 EKEVE
Sc_Est2	366 WLFRLIPKIIQTFEYCTEISSTVT.IVYFRHDTWNKLIPTFI	9 ENNV

Telomerase RT Motifs (Fingers)

	MOTIF 1	MOTIF 2	MOTIF A	MOTIF B'
TRT con	R iPKK	fr I	p lyF D	cYD i
hTRT	11 SRLRFIPKPDG 0 LRPIV	69 PELYFVKVDVTGAYDTI	104 YVQCQGIPOGSILSTLLCSLCY	Y q GIPQGS 15 1 Y
spTRT	10 AVIRLLPKKNT 0 FRLIT	66 RKYFVRIDIKSCYDRI	99 YLQKVGIPQGSILSSFLCHFYM	
Ea_p123	10 GKRLIPKKT 0 FRPIM	67 PKLFFATMDIEKCYDSV	117 YKQTKGIPQGLCVSSILSSFY	
Sc_Est2	13 SKMRIIPKKS 2 FRIIA	68 PELYFMKFDVKSCYDSI	85 YIREDLFGQSSLSAPIVDLVY	
RT con	p hh h K	hR h	h hDh AF h	hpQG pP hh h
			GY	

Telomerase RT Motifs (Palm, Primer Grip)

	MOTIF C	MOTIF D	MOTIF E
TRT con	l1lrl DDfL it	g	n K w g s l
hTRT	15 LLRLVDDFLLVT	15 GVPEYGCNVNLRKTVV	24 WCGLLLDTRTL 192
spTRT	16 VLLRVDDFLFIT	15 GFEKHNFTSLEKTVI	22 FFGFSVNMRL 176
Ea_p123	24 LLMLTDDYLLIT	15 VSRENGFKFNMKLQT	28 WIGISIDMKTL 174
Sc_Est2	18 LILKLADDFLIIS	15 GFQKYNKANRDKILA	25 WKHSSTMMNFH 141
RT con	h Y DDhhh	Gh h cK h	hLG h
	F		

FIG. 11

181 GGACCCGGCGGCTTTCCGCGCGCTGGTGGCCCAGTGCCTGGTGTGCGTGCCCTGGGACGC
CCTGGGCCGCCGAAAGGCGCGCGACCACCGGGTCACGGACCACACGCACGGGACCCTGCG

NFkB_CS1
GGGRQTYYYQC
NFkB-MHC-I.2
TGGGCTTCCCC

241 ACGGCCGCCCCCGCCGCCCCCTCCTTCCGCCAGGTGGGCCTCCCCGGGGTCGGCGTCCG
TGCCGGCGGGGGCGGCGGGGGAGGAAGGCGGTCCACCCGGAGGGGCCCCAGCCGCAGGC

Intron1

301 GCTGGGGTTGAGGGCGGCCGGGGGAACCAGCGACATGCGGAGAGCAGCGCAGGCGACTC
CGACCCCAACTCCCGCCGGCCCCCTTGGTCGCTGTACGCCTCTCGTCGCGTCCGCTGAG

NFkB_CS1
GGGRQTYYYQC
NFkB_CS2
RGGGRMTYYCC
Topo_II_cleavage_site
RNYNNCNGYNGKTNINY
*****>

361 AGGGCGCTTCCCCCGCAGGTGTCCTGCCTGAAGGAGCTGGTGGCCCGAGTGCTGCAGAGG
TCCCGCGAAGGGGGCGTCCACAGGACGGACTTCCTCGACCACCGGGCTCACGACGTCTCC

FIG. 12

1 AAAACCCCCAA AACCCCCAAAA CCCCTTTT TAG AGCCCTGCAG TTGGAAATAT
51 AACCTCAGTA TTAATAAGCT CAGATTTTAA ATATTAATTA CAAACCTAA
101 ATGGAGGTG ATGTTGATAA TCAAGCTGAT AATCATGGCA TTCACCTCAGC
151 TCTTAAGACT TGTGAAGAAA TTAAAGAAGC TAAAACGTTG TACTCTTGGA
201 TCCAGAAAGT TATTAGATGA AGAAATCAAT CTCAAAGTCA TTATAAAGAT
251 TTAGAAGATA TTAAATATTT TGCGCAGACA AATATTGTTG CTACTCCACG
301 AGACTATAAT GAAGAAGATT TTAAAGTTAT TGCAAGAAAA GAAGTATTTT
351 CAACTGGACT AATGATCGAA CTTATTGACA AATGCTTAGT TGAACCTCTT
401 TCATCAAGCG ATGTTTCAGA TAGACAAAAA CTTCAATGAT TTGGATTTCA
451 ACTTAAGGGA AATCAATTAG CAAAGACCCA TTTATTAACA GCTCTTTCAA
501 CTCAAAGCA GTATTTCTTT CAAGACGAAT GGAACCAAGT TAGAGCAATG
551 ATTGGAAATG AGCTCTTCCG ACATCTCTAC ACTAAATATTT TAATATTCCA
601 GCGAACTTCT GAAGGAATCT TTGTTCAATT TTGCGGGAAT AACGTTTTTG
651 ATCATTTGAA AGTCAACGAT AAGTTTGACA AAAAGCAAAA AGGTGGAGCA
701 GCAGACATGA ATGAACCTCG ATGTTGATCA ACCTGCAAAAT ACAATGTCAA
751 GAATGAGAAA GATCACTTTC TCAACAACAT CAACGTGCCG AATTGGAATA
801 ATATGAAATC AAGAACCAGA ATATTTTATTT GCACTCATTT TAATAGAAAT
851 AACCAATTCT TCAAAAAGCA TGAGTTTGTG AGTAACAAAA ACAATATTTT
901 AGCGATGGAC AGAGCTCAGA CGATATTCAC GAATATATTC AGATTTAATA
951 GAATTAGAAA GAAGCTAAAA GATAAGGTTA TCGAAAAAAT TGCCTACATG
1001 CTTGAGAAAG TCAAAGATTT TAACTTCAAC TACTATTTAA CAAAATCTTG
1051 TCCTCTTCCA GAAAATTGGC GGAACGGAA ACAAAAAATC GAAAACCTGA
1101 TAAATAAAAC TAGAGAAGAA AAGTCGAAGT ACTATGAAGA GCTGTTTAGC
1151 TACACAAC TG ATAATAAATG CGTCACACAA TTTATTAATG AATTTTTCTA
1201 CAATATACT CCCAAAGACT TTTTGACTGG AAGAAACCGT AAGAATTTTCT
1251 AAAAGAAAGT TAAGAAATAT GTGGAACATA ACAAGCATGA ACTCATTCAC
1301 AAAAATTTAT TGCTTGAGAA GATCAATACA AGAGAAATAT CATGGATGCA
1351 GGTGAGACC TCTGCAAAGC ATTTTATTA TTTTGATCAC GAAAACATCT
1401 ACGTCTTATG GAAATTGCTC CGATGGATAT TCGAGGATCT CGTCGTCTCG
1451 CTGATTAGAT GATTTTCTTA TGTCACCGAG CAACAGAAAA GTTACTCCAA
1501 AACCTATTAC TACAGAAAGA ATATTTGGGA CGTCATTATG AAAATGTCAA
1551 TCGCAGACTT AAAGAAGGAA ACGCTTGCTG AGGTCCAAGA AAAAGAGGTT
1601 GAAGAATGGA AAAAGTCGCT TGGATTTGCA CCTGGAAAAC TCAGACTAAT
1651 ACCGAAGAAA ACTACTTTCC GTCCAATTAT GACTTTCAAT AAGAAGATTG
1701 TAAATTCAGA CCGGAAGACT ACAAATTAAT CTACAAATAC GAAGTTATTG
1751 AACTCTCACT TAATGCTTAA GACATTGAAG AATAGAATGT TTAAAGATCC
1801 TTTTGGATTG GCTGTTTTTA ACTATGATGA TGTAATGAAA AAGTATGAGG
1851 AGTTTGTTTG CAAATGGAAG CAAGTTGGAC AACCAAAAC CTCTTTGCA
1901 ACTATGGATA TCGAAAAGTG ATATGATAGT GTAAACAGAG AAAAATATC
1951 AACATTCCTA AAAACTACTA AATTACTTTC TTCAGATTTC TGGATTATGA
2001 CTGCACAAAT TCTAAAGAGA AAGAATAACA TAGTTATCGA TTCGAAAAAC
2051 TTTAGAAAGA AAGAAATGAA AGATTATTTT AGACAGAAAT TCCAGAAGAT
2101 TGCACTTGAA GGAGGACAAT ATCCAACCTT ATTCAGTGTT CTTGAAAATG
2151 AACAAAATGA CTTAAATGCA AAGAAAACAT TAATTGTTGA AGCAAAGCAA
2201 AGAAATTATT TTAAGAAAGA TAACTTACTT CAACCAAGTCA TTAATATTTG
2251 CCAATATAAT TACATTAACT TTAATGGGAA GTTTTATAAA CAAACAAAAG
2301 GAATTCCCTA AGGTCTTTGA GTTTCATCAA TTTTGTATC ATTTTATTAT
2351 GCAACATTAG AGGAAAGCTC CTTAGGATTC CTTAGAGATG AATCAATGAA

FIG. 13

2401	CCCTGAAAAT	CCAAATGTTA	ATCTTCTAAT	GAGACTTACA	GATGACTATC
2451	TTTTGATTAC	AACTCAAGAG	AATAATGCAG	TATTGTTTAT	TGAGAACTT
2501	ATAAACGTAA	GTCGTGAAAA	TGGATTTAAA	TTCAATATGA	AGAAACTACA
2551	GACTAGTTTT	CCATTAAAGTC	CAAGCAAAT	TGCAAATAC	GGAATGGATA
2601	GTGTTGAGGA	GCAAAATATT	GTTCAAGATT	ACTGCGATTG	GATTGGCATC
2651	TCAATTGATA	TGAAAACCTCT	TGCTTTAATG	CCAAATATTA	ACTTGAGAAT
2701	AGAAGGAATT	CTGTGTACAC	TCAATCTAAA	CATGCAAACA	AAGAAAGCAT
2751	CAATGTGGCT	CAAGAAGAAA	CTAAAGTCGT	TTTTAATGAA	TAACATTACC
2801	CATTATTTTA	GAAAGACGAT	TACAACCGAA	GACTTTGCGA	ATAAACTCT
2851	CAACAAGTTA	TTTATATCAG	GCGGTTACAA	ATACATGCAA	TGAGCCAAAG
2901	AATACAAGGA	CCACTTTAAG	AAGAAGCTTAG	CTATGAGCAG	TATGATCGAC
2951	TTAGAGGTAT	CTAAATTTAT	ATACTCTGTA	ACCAGAGCAT	TCTTTAAATA
3001	CCTTGTTGTC	AATATTAAGG	ATACAATTTT	TGGAGAGGAG	CATTATCCAG
3051	ACTTTTTCCT	TAGCACACTG	AAGCACTTTA	TTGAAATATT	CAGCACAAAA
3101	AAGTACATTT	TCAACAGAGT	TTGCATGATC	CTCAAGGCAA	AAGAAGCAAA
3151	GCTAAAAAGT	GACCAATGTC	AATCTCTAAT	TCAATATGAT	GCATAGTCGA
3201	CTATTCTAAC	TTATTTTGGA	AAGTTAATTT	TCAATTTTGT	TCTTATATAC
3251	TGGGGTTTTG	GGGTTTTGGG	GTTTTGGGG		

FIG. 13
(Continued)

1	MEVDVDNQAD	NHGIHSALKT	CEEIKEAKTL	YSWIKQVIRC	RNQSQSHYKD
51	LEDIKIFAQT	NIVATPRDYN	EEDFKVIARK	EVFSTGLMIE	LIDKCLVELL
101	SSSDVSDRQK	LQCFGFQLKG	NGLAKTHLLT	ALSTQKQYFF	QDEWNQVRAM
151	IGNELFRHLY	TKYLIFQRTS	EGTLVQFCGN	NVFDHLKVND	KFDKKQKGA
201	ADMNEPRCCS	TCKYNVKNK	DHFLNNINVP	NWNMMKSRTR	IFYCTHFNRN
251	NQFFKKHEFV	SNKNNISAMD	RAQTIFTNIF	RFNRIRKKLK	DKVIEKIAYM
301	LEKVKDFNFN	YYLTKSCPLP	ENWRERKQKI	ENLINKTREE	KSKYYEELFS
351	YTDDNKCVTQ	FINEFFYNIL	PKDFLTGRNR	KNFQKKVKKY	VELNKHელი
401	KNLLEKINT	REISWMQVET	SAKHFFYFDH	ENIYVLWKLL	RWIFEDLVVS
451	LIRCFYFVTE	QQKSYSKTY	YRKNIVDVM	KMSIADLKKE	TLAEVQEKEV
501	EEWKSLGFA	PGKLRLIPK	TTFRPIMTFN	KKIVNSDRKT	TKLTTNTKLL
551	NSHMLMLTK	NRMFKDPFGF	AVFNYYDDVMK	KYEEFVCKWK	QVGQPKLFFA
601	TMDIEKCYDS	VNREKLSTFL	KTTKLLSSDF	WIMTAQILKR	KNNIVIDSKN
651	FRKKEMKDYF	RQKFQKIALE	GGQYPTLFSV	LENEQNDLNA	KKTLIVEAKQ
701	RNYFKKDNLL	QPVINICQYN	YINFNGKFYK	QTKGIPQGLC	VSSILSSFY
751	ATLEESSLGF	LRDESMNPEN	PNVNLLMRLT	DDYLLITTQE	NNAVLFIEKL
801	INVSRENGFK	FNMKKLQTSF	PLSPSKFAKY	GMDSVEEQNI	VQDYCDWIGI
851	SIDMKTLALM	PNINLRIEIG	LCTLNLMQ	KKASMWLKKK	LKSFLMNNIT
901	HYFRKTITTE	DFANKTLNKL	FISGGYKYM	CAKEYKDHF	KNLAMSSMID
951	LEVSKIISVS	TRAFFKYLVC	NIKDTIFGEE	HYPDFFLSTL	KHFIEIFSTK
1001	KYIFNRVCM	LKAKEAKLKS	DQCQSLIQYD	A	

FIG. 14

1 ggtaaccgatttacttcttcttccataaagctaattgcttccctgaacgctcctaaatctctggaataatttttacaaga 80
 81 actcaataacaataccaagtcacaaattccaatatgaagtggtattagtgatcgataataatttctattttatcggtcgtta 160
 161 ccaagtataaggacaaaagaacaaacttctccctcaagacttttacttttaataatttacttttcaatatatttcg 240
 241 ggttcgttactttttaaactggttactgttttagcttctagccaacccggtgttttaccctccatggatat 320
 321 agtcttgagtagctcacagaaaacttacaactctctgatgagactatattagattcattacacgctgcataattc 400
 401 ttaacatggagccttacaccttttagtcacgctcgatgatggagtatttggtatccatccaacggttgccttgaagaag 480
 481 gttgataattatttgcaaatcatgcttcccttagtggtggtgaatccgcgaagttttttagctgacacgcttagcatg 560
 561 attgagatattcaaaaatttctatccactacaactccttaacgcggttttatttttctatttctattctcatgtt 640
 641 ccaaatatgatcatctcgattaggttttttccggtttactcctggaaatcgtaacgttttccactattcccttaata 720
 721 ataactaaaattagtttcgcttataattgatatagtagtaagaagattggtgattctactcgtgtaattttattagtt 800
 801 gatactttgcaaaaacattttatttagctatcatatataaaaaaacctataataataataataataatttcgcggtc 880
 881 actatttttaaaacggttatgatcagtaggacacattgcatatatatagttatgcttaattggttacttgaacttgc 958

 959 ATG ACC GAA CAC CAT ACC CCC AAA AGC AGG ATT CTT CGC TTT CTA GAG AAT CAA TAT GTA 1018
 1 M T E H H T P K S R I L R F L E N Q Y V 20

 1019 TAC CTA TGT ACC TTA AAT GAT TAT GTA CAA CTT GTT TTG AGA GGG TCG CCG GCA AGC TCG 1078
 21 Y L C T L N D Y V Q L V L R G S P A S S 40

 1079 TAT AGC AAT ATA TGC GAA CGC TTG AGA AGC GAT GTA CAA ACG TCC TTT TCT ATT TTT CTT 1138
 41 Y S N I C E R L R S D V Q T S F S I F L 60

 1139 CAT TCG ACT GTA GTC GGC TTC GAC AGT AAG CCA GAT GAA GGT GTT CAA TTT TCT TCT CCA 1198
 61 H S T V V G F D S K P D E G V Q F S S P 80

 1199 AAA TGC TCA CAG TCA GAG gtatatatattttgtttgtattttttctattcgggatagctaataatgggcag 1272
 81 K C S Q S E 86

 1273 CTA ATA GCG AAT GTT GTA AAA CAG ATG TTC GAT GAA AGT TTT GAG CGT CGA AGG AAT CTA 1332
 87 L I A N V V K Q M F D E S F E R R N L 106

 1333 CTG ATG AAA GGG TTT TCC ATG gtaaggtattcttaattgtgaaaataattacctgcaattactgtttcaagaga 1405
 107 L M K G F S M 113

 1406 ttgtatttaaccgataaag AAT CAT GAA GAT TTT CGA GCC ATG CAT GTA AAC GGA GTA CAA AAT 1469
 114 N H E D F R A M H V N G V Q N 128

FIG. 15

1470 GAT CTC GTT TCT ACT TTT CCT AAT TAC CTT ATA TCT ATA CTT GAG TCA AAA AAT TGG CAA 1529
129 D L V S T F P N Y L I S I L E S K N W Q 148
1530 CTT TTG TTA GAA AT gtaataaccggttaagatgttgcgcactttgaacaagactgacaagtatag T ATC GGC 1601
149 L L L E I 155
1602 AGT GAT GCC ATG CAT TAC TTA TCC AAA GGA AGT ATT TTT GAG GCT CTT CCA AAT GAC 1661
156 S D A M H Y L L S K G S I F E A L P N D 175
1662 AAT TAC CTT CAG ATT TCT GGC ATA CCA CTT TTT AAA AAT AAT GTG TTT GAG GAA ACT GTG 1721
176 N Y L Q I S G I P L F K N N V F E E T V 195
1722 TCA AAA AAA AGA AAG CGA ACC ATT GAA ACA TCC ATT ACT CAA AAT AAA AGC GCC CGC AAA 1781
196 S K K R K R T I E T S I T Q N K S A R K 215
1782 GAA GTT TCC TGG AAT AGC ATT TCA ATT AGT AGG TTT AGC ATT TTT TAC AGG TCA TCC TAT 1841
216 E V S W N S I S I S R F S I F Y R S S Y 235
1842 AAG AAG TTT AAG CAA G gtaactaataactgttatccttcataactaatttttag AT CTA TAT TTT AAC 1907
236 K K F K Q D L Y F N 245
1908 TTA CAC TCT ATT TGT GAT CGG AAC ACA GTA CAC ATG TGG CTT CAA TGG ATT TTT CCA AGG 1967
246 L H S I C D R N T V H M W L Q W I F P R 265
1968 CAA TTT GGA CTT ATA AAC GCA TTT CAA GTG AAG CAA TTG CAC AAA GTG ATT CCA CTG GTA 2027
266 Q F G L I N A F Q V K Q L H K V I P L V 285
2028 TCA CAG AGT ACA GTT GTG CCC AAA CGT CTC CTA AAG GTA TAC CCT TTA ATT GAA CAA ACA 2087
286 S Q S T V V P K R L L K V Y P L I E Q T 305
2088 GCA AAG CGA CTC CAT CGT ATT TCT CTA TCA AAA GTT TAC AAC CAT TAT TGC CCA TAT ATT 2147
306 A K R L H R I S L S K V Y N H Y C P Y I 325
2148 GAC ACC CAC GAT GAT GAA AAA ATC CTT AGT TAT TCC TTA AAG CCG AAC CAG GTG TTT GCG 2207
326 D T H D D E K I L S Y S L K P N Q V F A 345
2208 TTT CTT CGA TCC ATT CTT GTT CGA GTG TTT CCT AAA TTA ATC TGG GGT AAC CAA AGG ATA 2267
346 F L R S I L V R V F P K L I W G N Q R I 365

FIG. 15
(CONTINUED)

2268 TTT GAG ATA ATA TTA AAA G gtattgtataaaatttattaccactaacgattttaccag AC CTC GAA ACT 2336
366 F E I I L K D L E T 375

2337 TTC TTG AAA TTA TCG AGA TAC GAG TCT TTT AGT TTA CAT TAT TTA ATG AGT AAC ATA AAG 2396
376 F L K L S R Y E S F S L H Y L M S N I K 395

2397 gtaatatgccaaaattttttaccatttaataacaatcag ATT TCA GAA ATT GAA TGG CTA GTC CTT GGA 2465
396 I S E I E W L V L G 405

2466 AAA AGG TCA AAT GCG AAA ATG TGC TTA AGT GAT TTT GAG AAA CGC AAG CAA ATA TTT GCG 2525
406 K R S N A K M C L S D F E K R K Q I F A 425

2526 GAA TTC ATC TAC TGG CTA TAC AAT TCG TTT ATA ATA CCT ATT TTA CAA TCT TTT TTT TAT 2585
426 E F I Y W L Y N S F I I P I L Q S F F Y 445

2586 ATC ACT GAA TCA AGT GAT TTA CGA AAT CGA ACT GTT TAT TTT AGA AAA GAT ATT TGG AAA 2645
446 I T E S S D L R N R T V Y F R K D I W K 465

2646 CTC TTG TGC CGA CCC TTT ATT ACA TCA ATG AAA ATG GAA GCG TTT GAA AAA ATA AAC GAG 2705
466 L L C R P F I T S M K M E A F E K I N E 485

2706 gtatttttaaaagtatttttttgcaaaaagctaataattttccag AAC AAT GTT AGG ATG GAT ACT CAG AAA ACT 2775
486 N N V R M D T Q K T 495

2776 ACT TTG CCT CCA GCA GTT ATT CGT CTA TTA CCT AAG AAG AAT ACC TTT CGT CTC ATT ACG 2835
496 T L P P A V I R L L P K K N T F R L I T 515

2836 AAT TTA AGA AAA AGA TTC TTA ATA AAG gtatttaatttttggtgcataatgtactttacttctaatttatta 2906
516 N L R K R F L I K 524

2907 ttagcag ATG GGT TCA AAC AAA AAA ATG TTA GTC AGT ACG AAC CAA ACT TTA CGA CCT GTG 2967
525 M G S N K K M L V S T N Q T L R P V 542

2968 GCA TCG ATA CTG AAA CAT TTA ATC AAT GAA GAA AGT AGT GGT ATT CCA TTT AAC TTG GAG 3027
543 A S I L K H L I N E S S G I P F N L E 562

3028 GTT TAC ATG AAG CTT CTT ACT TTT AAG AAG GAT CTT CTT AAG CAC CGA ATG TTT GG gtaat 3088
563 V Y M K L L T F K K D L L K H R M F G 581

FIG. 15
(CONTINUED)

17/34

3089 tatataatgcgcgattccctcattattatttcag G CGT AAG AAG TAT TTT GTA CGG ATA GAT ATA 3155
 582 R K K Y F V R I D I
 3156 AAA TCC TGT TAT GAT CGA ATA AAG CAA GAT TTG ATG TTT CGG ATT GTT AAA AAG AAA CTC 3215
 592 K S C Y D R I K Q D L M F R I V K K L
 3216 AAG GAT CCC GAA TTT GTA ATT CGA AAG TAT GCA ACC ATA CAT GCA ACA AGT GAC CGA GCT 3275
 612 K D P E F V I R K Y A T I H A T S D R A
 3276 ACA AAA AAC TTT GTT AGT GAG GCG TTT TCC TAT T gtaagttatttttttcattggaatttttaacaa 3343
 632 T K N F V S E A F S Y F
 3344 attcttttttag TT GAT ATG GTG CCT TTT GAA AAA GTC GTG CAG TTA CTT TCT ATG AAA ACA 3405
 644 D M V P F E K V V Q L L S M K T
 3406 TCA GAT ACT TTG TTT GTT GAT TTT GTG GAT TAT TGG ACC AAA AGT TCT TCT GAA ATT TTT 3465
 660 S D T L F V D F V D Y W T K S S E I F
 3466 AAA ATG CTC AAG GAA CAT CTC TCT GGA CAC ATT GTT AAG gtataccaattgtgaattgtaataaca 3532
 680 K M L K E H L S G H I V K
 3533 ctaatgaaactag ATA GGA AAT TCT CAA TAC CTT CAA AAA GTT GGT ATC CCT CAG GGC TCA 3593
 693 I G N S Q Y L Q K V G I P Q G S
 3594 ATT CTG TCA TCT TTT TTG TGT CAT TTC TAT ATG GAA GAT TTG ATT GAT GAA TAC CTA TCG 3653
 709 I L S S F L C H F Y M E D L I D E Y L S
 3654 TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA GTA GTC GAC GAT TTC CTC TTT ATA ACA 3713
 729 F T K K G S V L L R V V D D F L F I T
 3714 GTT AAT AAA AAG GAT GCA AAA AAA TTT TTG AAT TTA TCT TTA AGA G gtgagttgctgtcattcc 3777
 749 V N K K D A K K F L N L S L R G
 3778 taagttctaaccgttgaag GA TTT GAG AAA CAC AAT TTT TCT ACG AGC CTG GAG AAA ACA GTA 3840
 765 F E K H N F S T S L E K T V
 3841 ATA AAC TTT GAA AAT AGT AAT ATA AAC AAT ACT TTT TTT AAT GAA AGC AAG AAA 3900
 779 I N F E N S N G I I N N T F F N E S K K
 798

FIG. 15
(CONTINUED)

18/34

3901 AGA ATG CCA TTC TTC GGT TTC TCT GTG AAC ATG AGG TCT CTT GAT ACA TTG TTA GCA TGT 3960
 799 R M P F F G F S V N M R S L D T L L A C 818
 3961 CCT AAA ATT GAT GAA GCC TTA TTT AAC TCT ACA TCT GTA GAG CTG ACG AAA CAT ATG GGG 4020
 819 P K I D E A L F N S T S V E L T K H M G 838
 4021 AAA TCT TTT TTT TAC AAA ATT CTA AG gtatactgtgtaactgaataatagctgacaaaataatcag A TCG 4089
 839 K S F F Y K I L R S 848
 4090 AGC CTT GCA TCC TTT GCA CAA GTA TTT ATT GAC ATT ACC CAC AAT TCA AAA TTC AAT TCT 4149
 849 S L A S F A Q V F I D I T H N S K F N S 868
 4150 TGC TGC AAT ATA TAT AGG CTA GGA TAC TCT ATG TGT ATG AGA GCA CAA GCA TAC TTA AAA 4209
 869 C C N I Y R L G Y S M C M R A Q A Y L K 888
 4210 AGG ATG AAG GAT ATA TTT ATT CCC CAA AGA ATG TTC ATA ACG G gtgagtacttattttaaactaga 4274
 889 R M K D I F I P Q R M F I T D 903
 4275 aaagtcattaattaaccttag AT CTT TTG AAT GTT ATT GGA AGA AAA ATT TGG AAA AAG TTG GCC 4339
 904 L L N V I G R K I W K K L A 917
 4340 GAA ATA TTA GGA TAT ACG AGT AGG CGT TPC TTG TCC TCT GCA GAA GTC AAA TG gtacgtgtc 4401
 918 E I L G Y T S R R F L S S A E V K W 935
 4402 ggtctcgagacttcagcaaatattgacacatcag G CTT TTT TGT CTT GGA ATG AGA GAT GGT TTG AAA 4468
 936 L F C L G M R D G L K 946
 4469 CCC TCT TTC AAA TAT CAT CCA TGC TTC GAA CAG CTA ATA TAC CAA TTT CAG TCA TTG ACT 4528
 947 P S F K Y H P C F E Q L I Y Q F Q S L T 966
 4529 GAT CTT ATC AAG CCG CTA AGA CCA GTT TTG CGA CAG GTG TTA TTT TTA CAT AGA ATA 4588
 967 D L I K P L R P V L R Q V L F L H R R I 986
 4589 GCT GAT TAA tgcatttttcaatttattatatacatcctttattactggtgtctttaaacaaatattattactaagtata 4665
 987 A D * 989

FIG. 15
(CONTINUED)

19/34

4666 gctgacccccaaagcaatactataggatttctagtaaaagtaaaataatcctcgttattagtttttgattgacttgtct 4745
4746 ttatccttataacttttaagaaagattgacagtgttgctgactactgcccacatgcccattaaacgggagtggttaaaca 4825
4826 ttaaaagtaatacatgaggctaatactcctttcatttagaataaggaaagtgtgtttctataatgaataaatgccgcacta 4905
4906 atgcaaaaagacgaagattatcttctaacaagggttlaagcatatccgaaggaaagagagataatatacccagtggt 4985
4986 gttgaagaaagcaaggataatttggaacaagcttctgcagatgacaggctaaaattttggtgaccgaaattttggtaaaagc 5065
5066 ccaggttatccatggtggccggccttgctactgagacgaaagaaactaaggatagtttgaataactaataagctcattta 5145
5146 atgtcttataaagggttttggttttttccctgacttcaatttgcatgggtgaaagaaatagtttaagccattattggat 5225
5226 tcgaaaatagccaaatttcttggttccctcaaggcggaagtctaaagaacttattgaagcttatgaggcttcaaaaactcc 5305
5306 tcctgatttaaggagggaatcttccaccgatgaggaatggatagcttatacagctgctgaggagaagcctaattttttgc 5385
5386 aaaaaagaaaatacatbgggagacatctcttgatgaatcagatcgaggagatctccagcggtccttgatgtcaata 5465
5466 acttctatttctgaaatgtatgggtcctactgtcggttcgacttctcgtagctctacgcaggttaagtgaacaaaggtacc 5544

FIG. 15
(CONTINUED)

20/34

1	gcagcgctgc	gtcctgctgc	gcacgtggga	agccctggcc	ccggccaccc	ccgcgatgcc
61	gcgcgctccc	cgctgcccag	ccgtgcgctc	cctgctgcgc	agccactacc	gcgaggtgct
121	gccgctggcc	acgttccgtgc	ggcgccctggg	gccccagggc	tggcggctgg	tgcagcgccg
181	ggaccgccgg	gctttccgcg	cgctgggtggc	ccagtgcctg	gtgtgcgtgc	cctggggacgc
241	acggccgcgcc	cccgccgcgcc	cctccttccg	ccaggtgtcc	tgccctgaagg	agctgggtggc
301	ccgagtgctg	cagaggctgt	gcgagcgccg	cgcaagaac	gtgctggcct	tcggcttcgc
361	gctgctggac	ggggcccgcg	ggggccccc	cgaggccttc	accaccagcg	tgcgcagcta
421	cctgccccaac	acggtgaccg	acgcactgcg	ggggagcggg	gcgtgggggc	tgctgctgcg
481	ccgcgtgggc	gacgacgtgc	tggttcacct	gctggcacgc	tgccgcgtct	ttgtgctggt
541	ggctcccagc	tgccgctacc	aggtgtgcgg	gccgcccgtg	taccagctcg	gcgctgccac
601	tcaggcccg	cccccgccac	acgctagtgg	accccgagg	cgtctgggat	gcgaacgggc
661	ctgggaacct	agcgtcaggg	agggcgggg	ccccctgggc	ctgccagccc	cggtgtgcgag
721	gcagcccg	ggcgagcgga	gcccgaagtct	gcccgttgccc	aagaggcca	ggcgtggcgc
781	tgcccctgag	ccggagcgga	cgcccgttgg	gcaggggtcc	tgggcccacc	cggtgtggac
841	gcgtggaccg	agtgaccgtg	gtttctgtgt	ggtgtcacct	gccagaccgg	ccgaagaagc
901	cacctctttg	gaggggtgcg	tctctggcac	gcgccactcc	cacctatccg	tgggcccgcga
961	gcaccacgg	ggccccccat	ccacactcgc	gccaccacgt	ccctgggaca	cgccctgtcc
1021	cccggtgtac	ggcgagacca	agcacttctc	ctactcctca	ggcgacaagg	agcagctgcg
1081	gcccctccttc	ctactcagct	ctctgaggcc	cagcctgact	ggcgctcgga	ggctcgtgga
1141	gaccatcttt	ctgggttcca	ggccctggat	gccagggact	ccccgcaggt	tgccccgcct
1201	gcccacgcgc	tactggcaaa	tgccgcccct	gtttctggag	ctgcttggga	accacgcgca
1261	gtgcccctac	ggggtgctcc	tcaagacgca	ctgcccgcgtg	cgagctgcgg	tcaccccagc
1321	agccggtgtc	tgtgcccggg	agaagcccca	gggctctgtg	gcggcccccg	aggaggagga
1381	cacagacccc	cgctgcctgg	tgcagctgct	ccgccagcac	agcagcccct	ggcaggtgta
1441	cggtctcgtg	cgggcctggc	tgccgcggct	gggtgccccca	ggcctctggg	gctccaggca
1501	caacgaacgc	cgcttctctc	ggaacaccaa	gaagtccatc	tccttgggga	agcatgccaa
1561	gctctcgtcg	caggagctga	cgtggaagat	gagcgtgcgg	gactgcgctt	ggctgcgcag
1621	gagcccagg	gttggtctgt	ttccggccgc	agagcacctg	ctgcgtgagg	agatcctggc
1681	caagtctcgt	cactggctga	tgagtgtgta	cgctcgtcag	ctgctcaggt	ctttctttta
1741	tgtcacggag	accacgtttc	aaaagaacag	gctctttttc	taccggaaga	gtgtctggag
1801	caagttgcaa	agcattggaa	tcagacagca	cttgaagagg	gtgcagctgc	gggagctgtc
1861	ggaagcagag	gtcaggcagc	atcgggaagc	caggcccgcg	ctgctgacgt	ccagactccg
1921	cttcaccccc	aagcctgacg	ggctgcggcc	gattgtgaac	atggactacg	tcggtgggagc
1981	cagaacgttc	cgcagagaaa	agagggccga	gcgtctcacc	tcgagggtga	aggcactggt
2041	cagcgtgctc	aactacgagc	gggcccggcg	ccccggccctc	ctgggcccct	ctgtgctggg
2101	cctggacgat	atccacaggg	cctggcgcac	cttcgtgctg	cgtgtgcggg	cccaggaccc
2161	ccgccttgag	ctgtactttg	tcaagttgga	tgtgacgggc	gcgtacgaca	ccatccccca
2221	ggacaggctc	acggaggtca	tcgccagcat	catcaaacc	cagaacacgt	actgcgtgcg
2281	tcggtatgcc	gtgggtccaga	aggccgcccc	tgggcacgtc	cgcaaggcct	tcaagagcca
2341	cgctcttacc	ttgacagacc	tcagccgta	catgcgacag	ttcgtggctc	acctgcagga
2401	gaccagcccg	ctgagggtag	ccgtcgtcat	cgagcagagc	tcctccctga	atgaggccag
2461	cagtggcctc	ttcgacgtct	tcctacgctt	catgtgccac	cacgccgtgc	gcacagggg
2521	caagtccctac	gtccagtgc	aggggatccc	gcagggctcc	atcctctcca	cgctgctctg
2581	cagcctgtgc	tacggcgaca	tgagaaacaa	gctgtttgcg	gggattccgc	gggacgggct
2641	gctcctgcgt	ttggtggatg	atctcttgtt	ggtgacacct	cacctcaacc	acgcgaaaac
2701	cttcctcagg	accctggctc	gaggtgtccc	tgagtatggc	tgctgtgtga	acttgccgaa
2761	gacagtgggtg	aacttccctg	tagaagacga	ggccctgggt	ggcacggctt	ttgttcagat
2821	gccggccccc	ggcctattcc	cctggtgcgg	cctgctgctg	gatacccgga	ccctggaggt
2881	gcagagcgac	tactccagct	atgcccggac	ctccatcaga	gccagtctca	ccttcaaccg
2941	cggtctcaag	gctgggagga	acatgcgtcg	caaactcttt	ggggtcttgc	ggctgaagtg
3001	tcacagcctg	tttctggatt	tgcaggtgaa	cagcctccag	acggtgtgca	ccaacatcta
3061	caagatcctc	ctgctgcagg	cgtaacaggt	tcacgcatgt	gtgctgcagc	tccattttca
3121	tcagcaagtt	tggaagaacc	ccacattttt	cctgcgcgtc	atctctgaca	cgccctccct
3181	ctgctactcc	atcctgaaag	ccaagaacgc	agggatgtcg	ctgggggcca	agggcgccgc
3241	cgccctctctg	ccctccgagg	ccgtgcagtg	gctgtgccac	caagcattcc	tgctcaagct
3301	gactcgacac	cgtgtcacct	acgtgccact	cctggggtca	ctcaggacag	cccagacgca
3361	gctgagtcgg	aagctcccgg	ggacgacgct	gactgccctg	gaggccgcag	ccaacccggc
3421	actgccctca	gacttcaaga	ccatccctgga	ctgatggcca	cccggccaca	gccaggccga
3481	gagcagacac	cagcagccct	gtcacgcggg	gctctacgtc	ccagggaggg	agggggcgcc
3541	cacacccagg	cccgacccgc	tgaggtctgt	agggcctgag	gagtggttgg	ccgagggcctg
3601	catgtccggc	tgaaggctga	gtgtccggct	gaggcctgag	cgagtgtcca	gccaaaggct
3661	gagtgtccag	cacacctgcc	gtcttcaact	ccccacaggc	tggcgctcgg	ctccacccca
3721	gggcccagctt	ttcctcacca	ggagcccggc	ttccactccc	cacataggaa	tagtccatcc
3781	ccagattcgc	catgtttcac	cctctgcctc	gcctcctttt	gccttccacc	cccaccatcc
3841	aggtggagac	cctgagaagg	accctgggag	ctctgggaat	ttggagtgc	caaagggtgtg
3901	ccctgtacac	aggcgaggac	cctgcacctg	gatgggggtc	cctgtgggtc	aaattggggg
3961	gaggtgctgt	gggagtaaaa	tactgaatat	atgagttttt	cagttttttaa	aaaaa

FIG. 16

MPRAPRCRAVRSLLRSHYREVLPLATFVRRRLGPQGWRLVQRGDP
AAFRALVAQCLVCPWDARPPPAAPSFRQVSCLKELVARVLQRL
CERGAKNVLAFGFALLDGARGGPPEAFTTSVRSYLPNTVTDALR
GSGAWGLLLRRVGDDVLVHLLARCALFVLVAPSCAYQVCGPPLY
QLGAATQARPPPHASGPRRRLGCERAWNHSVREAGVPLGLPAPG
ARRRGGSASRSLPLPKRPRRGAAPERTPVGQGSWAHPGRTRG
PSDRGFCVVSPARPAEEATSLEGALSGTRHSHPSVGRQHHAGPP
STSRPPRPWDTPCPPVYAETKHFYSSGDKEQLRPSFLLSSLRP
SLTGARRLVETIFLGSRPWMPGTPRRLPRLPQRYWQMRPLFLEL
LGNHAQCPYGVLLKTHCPLRAAVTPAAGVCAREKPGQSVAAPPEE
EDTDPRRLVQLLRQHSSPWQVYGFVRACLRRLVPPGLWGSRHNE
RRFLRNTKKFIISLGKHAKLSLQELTWKMSVRDCAWLRRSPGVGC
VPAAEHRLREEILAKFLHWMMSVYVVELLRSFFYVTETTFQKNR
LFFYRKSVWSKLSIGIRQHLLKRVQLRELSAEVRQHREARPAL
LTSRLRFIPKPDGLRPIVNM DYVVGARTFRREKRAERLTSRVKA
LFSVLNLYERARRPGLLGASVLGLDDIHRWRFTFVLRVRAQDPPP
ELYFVKVDVTGAYDTIPQDRLTEVIASIIKPQNTYCVRRYAVVQ
KAAHGHRVRAKAFKSHVSTLTDLQPYMRQFVAHLQETSPLRDAVVI
EQSSSLNEASSGLFDVFLRFMCHHAVRIRGKSYVQCQGIPOGSI
LTLCLSLCYGDMENKLFAGIRRDGLLLRLVDDFLVTPHLTHA
KTFLRTLVRGVPEYGCVVNLRKTVVNFVVEDEALGGTAFVQMPA
HGLFPWCGLLLDTRTLEVQSDYSSYARTSIRASLTFNRGFKAGR
NMRRKLFGLVRLKCHSLFLDLQVNSLQTVCTNIYKILLQAYRF
HACVLQLPFHQVWKNPTFFLRVISDTASLCYSILKAKNAGMSL
GAKGAAGPLPSEAVQWLCHQAFLLKLTRHRVTYVPLLGLSLRTAQ
TQLSRKLPGTTLTALEAAANPALPSDFKTILD

FIG. 17

GGCCAAGTTCTCTGCACTGGCTGATGAGTGTGTACGTCGTCGAGCTGCTCAGGTCTTTCTT
TTATGTACACGGAGACCACGTTTCAAAGAAGACAGGCTCTTTTCTACCGGAAGAGTGTCTG
GAGCAAGTTGCAAAGCATTTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCGGGAGCT
GTCGGAAGCAGAGGTGACGAGCATCGGGAAGCCAGGCCCGCCCTGCTGACGTCCAGACT
CCGCTTCATCCCCAAGCCTGACGGGCTGCGGCCGATTGTGAACATGGACTACGTCTGCGG
AGCCAGAAGCTTCCGAGAGAAAAGAGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACT
GTTACGCGTGTCAACTACGAGCGGGCGCGGCCCGCCCTCTGGGCGCCTCTGTGCT
GGGCTTGGACGATATCCACAGGGCCTGGCGCACCTTCGTGCTGCGTGTGCGGGCCAGGA
CCCGCCGCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGCGTACGACACCATCCC
CCAGGACAGGCTCACGAGGTTCATCGCCAGCATCATCAAACCCAGAACACGTACTGCGT
GCGTCCGTATGCCGTGGTCCAGAAGGCCGCCCATGGGCACGTCCGCAAGGCCCTCAAGAG
CCACGTCCTACGTCCAGTGCCAGGGGATCCCGCAGGGCTCCATCCTCTCCACGTGCTCT
GCAGCCTGTGCTACGGCGACATGGAGAACAAGCTGTTTGGCGGGATTGCGCGGGACGGC
TGCTCCTGCGTGTGGTGGATGATTTCTTGTGGTGACACCTCACCTCACCCACGCGAAAA
CCTTCCTCAGGACCCTGGTCCGAGGTGTCCCTGAGTATGGCTGCGTGGTGAACCTTGCAGGA
AGACAGTGGTGAACCTTCCCTGTAGAAGACGAGGCCCTGGGTGGCACGGCTTTTGTTCAGA
TGCCGGCCACGGCCTATTCCTTGGTGCGCCTGCTGCTGGATACCCGACCCCTGGAGG
TGCAGAGCGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCCAGTCTCACCTTCAACC
GCGGCTTCAAGGCTGGGAGGAACATGCGTCCGAAACTCTTTGGGGTCTTGCAGGCTGAAGT
GTCACAGCCTGTTTCTGGATTTGCAGGTGAACAGCCTCCAGACGGTGTGCACCAACATCT
ACAAGATCCTCCTGCTGCAGGCGTACAGGTTTACGCATGTGTGCTGCAGCTCCCATTTC
ATCAGCAAGTTTGAAGAACCACATTTTCTGCGCGTCATCTCTGACACGGCCTCCC
TCTGCTACTCCATCCTGAAAGCCAAGAAGCAGGGATGTGCTGGGGGCCAAGGGCGCCG
CCGGCC7TCTGCCCTCCGAGGCCGTGCACTGGCTGTGCCACCAAGCATTCCTGCTCAAGC
TGACTCGACACCGTGTACCTACGTGCCACTCCTGGGGTCACTCAGGACAGCCAGACGC
AGCTGAGTCGGAAGCTCCCGGGACGACGTGACTGCCCTGGAGGCCGAGCCAACCCGG
CACTGCCCTCAGACTTCAAGACCATCCTGGACTGATGGCCACCCGCCACAGCCAGGCCG
AGAGCAGACACCAGCAGCCCTGTACGCGCGGCTCTACGTCCCAGGGAGGGAGGGCGGC
CCACACCCAGGCCGTGACCCGTGGGAGTCTGAGGCCTGAGTGAGTGTGGCCGAGGCCCT
GCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGCCTGAGCGAGTGTCCAGCCAAGGGC
TGAGTGTCAGCACACCTGCGCTCTCACTTCCCCACAGGCTGGCGCTCGGCTCCACCCC
AGGGCCAGCTTTTCTCACCAGGAGCCCGCTTCCACTCCCCACATAGGAATAGTCCATC
CCCAGATTGCCATTGTTACCCCTCGCCCTGCCCTCCTTTGCCTTCCACCCCCACCATC
CAGGTGGAGACCTGAGAAGGACCCTGGGAGCTCTGGGAATTTGGAGTGACCAAAGGTGT
GCCCTGTACACAGGCGAGGACCCTGCACCTGGATGGGGTCCCTGTGGGTCAAATTGGGG
GGAGGTGCTGTGGGAGTAAAAATACTGAATATAGATTTTTTCAGTTTTTG0AAAAAAAAA
AAAAAAAAAAAAAAAAA

FIG. 18

FIG. 19

FIG. 20

```

                                130
tyr leu pro asn thr val thr asp ala leu arg gly ser gly ala
TAC CTG CCC AAC ACG GTG ACC GAC GCA CTG CGG GGG AGC GGG GCG

                                140
trp gly leu leu leu arg arg val gly asp asp val leu val his
TGG GGG CTG CTG CTG CGC CGC GTG GGC GAC GAC GTG CTG GTT CAC

                                150

                                160
leu leu ala arg cys ala leu phe val leu val ala pro ser cys
CTG CTG GCA CGC TGC GCG CTC TTT GTG CTG GTG GCT CCC AGC TGC

                                170
ala tyr gln val cys gly pro pro leu tyr gln leu gly ala ala
GCC TAC CAG GTG TGC GGG CCG CCG CTG TAC CAG CTC GGC GCT GCC

                                180

                                190
thr gln ala arg pro pro pro his ala ser gly pro arg arg arg
ACT CAG GCC CGG CCC CCG CCA CAC GCT AGT GGA CCC CGA AGG CGT

                                200
leu gly cys glu arg ala trp asn his ser val arg glu ala gly
CTG GGA TGC GAA CGG GCC TGG AAC CAT AGC GTC AGG GAG GCC GGG

                                210

                                220
val pro leu gly leu pro ala pro gly ala arg arg arg gly gly
GTC CCC CTG GGC CTG CCA GCC CCG GGT GCG AGG AGG CGC GGG GGC

                                230
ser ala ser arg ser leu pro leu pro lys arg pro arg arg gly
AGT GCC AGC CGA AGT CTG CCG TTG CCC AAG AGG CCC AGG CGT GGC

                                240

                                250
ala ala pro glu pro glu arg thr pro val gly gln gly ser trp
GCT GCC CCT GAG CCG GAG CGG ACG CCC GTT GGG CAG GGG TCC TGG

                                260
ala his pro gly arg thr arg gly pro ser asp arg gly phe cys
GCC CAC CCG GGC AGG ACG CGT GGA CCG AGT GAC CGT GGT TTC TGT

                                270

                                280
val val ser pro ala arg pro ala glu glu ala thr ser leu glu
GTG GTG TCA CCT GCC AGA CCC GCC GAA GAA GCC ACC TCT TTG GAG

                                290
gly ala leu ser gly thr arg his ser his pro ser val gly arg
GGT GCG CTC TCT GGC ACG CGC CAC TCC CAC CCA TCC GTG GGC CGC

                                300

                                310
gln his his ala gly pro pro ser thr ser arg pro pro arg pro
CAG CAC CAC GCG GGC CCC CCA TCC ACA TCG CGG CCA CCA CGT CCC

                                320
trp asp thr pro cys pro pro val tyr ala glu thr lys his phe
TGG GAC ACG CCT TGT CCC CCG GTG TAC GCC GAG ACC AAG CAC TTC

                                330

```

FIG. 20
(CONTINUED)

```

                                340
leu tyr ser ser gly asp lys glu gln leu arg pro ser phe leu
CTC TAC TCC TCA GGC GAC AAG GAG CAG CTG CGG CCC TCC TTC CTA

                                350
leu ser ser leu arg pro ser leu thr gly ala arg arg leu val
CTC AGC TCT CTG AGG CCC AGC CTG ACT GGC GCT CGG AGG CTC GTG

                                360

                                370
glu thr ile phe leu gly ser arg pro trp met pro gly thr pro
GAG ACC ATC TTT CTG GGT TCC AGG CCC TGG ATG CCA GGG ACT CCC

                                380
arg arg leu pro arg leu pro gln arg tyr trp gln met arg pro
CGC AGG TTG CCC CGC CTG CCC CAG CGC TAC TGG CAA ATG CGG CCC

                                390

                                400
leu phe leu glu leu leu gly asn his ala gln cys pro tyr gly
CTG TTT CTG GAG CTG CTT GGG AAC CAC GCG CAG TGC CCC TAC GGG

                                410
val leu leu lys thr his cys pro leu arg ala ala val thr pro
GTG CTC CTC AAG ACG CAC TGC CCG CTG CGA GCT GCG GTC ACC CCA

                                420

                                430
ala ala gly val cys ala arg glu lys pro gln gly ser val ala
GCA GCC GGT GTC TGT GCC CGG GAG AAG CCC CAG GGC TCT GTG GCG

                                440
ala pro glu glu glu asp thr asp pro arg arg leu val gln leu
GCC CCC GAG GAG GAG GAC ACA GAC CCC CGT CGC CTG GTG CAG CTG

                                450

                                460
leu arg gln his ser ser pro trp gln val tyr gly phe val arg
CTC CGC CAG CAC AGC AGC CCC TGG CAG GTG TAC GGC TTC GTG CGG

                                470
ala cys leu arg arg leu val pro pro gly leu trp gly ser arg
GCC TGC CTG CGC CGG CTG GTG CCC CCA GGC CTC TGG GGC TCC AGG

                                480

                                490
his asn glu arg arg phe leu arg asn thr lys lys phe ile ser
CAC AAC GAA CGC CGC TTC CTC AGG AAC ACC AAG AAG TTC ATC TCC

                                500
leu gly lys his ala lys leu ser leu gln glu leu thr trp lys
CTG GGG AAG CAT GCC AAG CTC TCG CTG CAG GAG CTG ACG TGG AAG

                                510

                                520
met ser val arg asp cys ala trp leu arg arg ser pro gly val
ATG AGC GTG CGG GAC TGC GCT TGG CTG CGC AGG AGC CCA GGG GTT

                                530
gly cys val pro ala ala glu his arg leu arg glu glu ile leu
GGC TGT GTT CCG GCC GCA GAG CAC CGT CTG CGT GAG GAG ATC CTG

                                540

```

FIG. 20
(CONTINUED)

550
ala lys phe leu his trp leu met ser val tyr val val glu leu
GCC AAG TTC CTG CAC TGG CTG ATG AGT GTG TAC GTC GTC GAG CTG

560 570
leu arg ser phe phe tyr val thr glu thr thr phe gln lys asn
CTC AGG TCT TTC TTT TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC

580
arg leu phe phe tyr arg lys ser val trp ser lys leu gln ser
AGG CTC TTT TTC TAC CGG AAG AGT GTC TGG AGC AAG TTG CAA AGC

590 600
ile gly ile arg gln his leu lys arg val gln leu arg glu leu
ATT GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG

610
ser glu ala glu val arg gln his arg glu ala arg pro ala leu
TCG GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG

620 630
leu thr ser arg leu arg phe ile pro lys pro asp gly leu arg
CTG ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CGG

640
pro ile val asn met asp tyr val val gly ala arg thr phe arg
CCG ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC

650 660
arg glu lys arg ala glu arg leu thr ser arg val lys ala leu
AGA GAA AAG AGG GCC GAG CGT CTC ACC TCG AGG GTG AAG GCA CTG

670
phe ser val leu asn tyr glu arg ala arg arg pro gly leu leu
TTC AGC GTG CTC AAC TAC GAG CGG GCG CGG CGC CCC GGC CTC CTG

680 690
gly ala ser val leu gly leu asp asp ile his arg ala trp arg
GGC GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC TGG CGC

700
thr phe val leu arg val arg ala gln asp pro pro pro glu leu
ACC TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCG CCT GAG CTG

710 720
tyr phe val lys val asp val thr gly ala tyr asp thr ile pro
TAC TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC

730
gln asp arg leu thr glu val ile ala ser ile ile lys pro gln
CAG GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC ATC AAA CCC CAG

740 750
asn thr tyr cys val arg arg tyr ala val val gln lys ala ala
AAC ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC

FIG. 20
(CONTINUED)

760
his gly his val arg lys ala phe lys ser his val leu arg pro
CAT GGG CAC GTC CGC AAG GCC TTC AAG AGC CAC GTC CTA CGT CCA

770 780
val pro gly asp pro ala gly leu his pro leu his ala ala leu
GTG CCA GGG GAT CCC GCA GGG CTC CAT CCT CTC CAC GCT GCT CTG

790
gln pro val leu arg arg his gly glu gln ala val cys gly asp
CAG CCT GTG CTA CGG CGA CAT GGA GAA CAA GCT GTT TGC GGG GAT

800 807
ser ala gly arg ala ala pro ala phe gly gly OP
TCG GCG GGA CGG GCT GCT CCT GCG TTT GGT GGA TGA TTTCTTGTGTTGGT

GACACCTCACCTCACCCACGCGAAAACCTTCCTCAGGACCCTGGTCCGAGGTGTCCCTGA
GTATGGCTGCGTGGTGAACCTTGCGGAAGACAGTGGTGAACCTTCCCTGTAGAAGACGAGGC
CCTGGGTGGCAGGCTTTTGTTCAGATGCCGGCCACGGCCTATTCCTTGGTGCAGGCT
GCTGCTGGATACCCGGACCTGGAGGTGCAGAGCGACTACTCCAGCTATGCCCCGACCTC
CATCAGAGCCAGTCTCACCTTCAACCGCGGCTTCAAGGCTGGGAGGAACATGCGTCGCAA
ACTCTTTGGGGTCTTGCGGTGAAGTGTACAGCCTGTTTCTGGATTTGCAGGTGAACAG
CCTCCAGACGGTGTGCACCAACATCTACAAGATCCTCCTGCTGCAGGCGTACAGGTTTCA
CGCATGTGTGCTGCAGCTCCATTTCATCAGCAAGTTTGAAGAACCCACATTTTTCCT
GCGCGTCATCTCTGACACGGCCTCCCTCTGCTACTCCATCCTGAAAGCCAAGAACGCAGG
GATGTGCTGCTGGGGGCCAAGGGCGCGCCGCGCCTCTGCCCTCCGAGGCCGTGCAGTGGCT
GTGCCACCAAGCATTCCTGCTCAAGCTGACTCGACACCGTGTACCTACGTGCCACTCCT
GGGGTCACTCAGGACAGCCCAGACGCAGCTGAGTCGGAAGCTCCCGGGGACGACGCTGAC
TGCCCTGGAGGCCGAGCCAACCCGGCACTGCCCTCAGACTTCAAGACCATCCTGGACTG
ATGGCCACCCGCCCACAGCCAGGCCGAGAGCAGACACCAGCAGCCCTGTCACGCCGGGCT
CTACGTCCCAGGGAGGGAGGGGCGGCCCACACCCAGGCCCGCACCGCTGGGAGTCTGAGG
CCTGAGTGAGTGTGTTGGCCGAGGCTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAG
GCCTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTCCAGCACACCTGCCGTCTTCACTTCCC
CACAGGCTGGCGCTCGGCTCCACCCAGGGCCAGCTTTTCCTCACCAGGAGCCCGGCTTC
CACTCCCCACATAGGAATAGTCCATCCCCAGATTTCGCCATTGTTACCCCTCGCCCTGCC
CTCCTTTGCCTTCCACCCCCACCATCCAGGTGGAGACCCTGAGAAGGACCCTGGGAGCTC
TGGGAATTTGGAGTGACCAAAGGTGTGCCCTGTACACAGGCGAGGACCCTGCACCTGGAT
GGGGTCCCTGTGGGTCAAATTGGGGGAGGTGCTGTGGGAGTAAAATACTGAATATATG
AGTTTTTCAGTTTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIG. 20
(CONTINUED)

[illegible][illegible]

4321 GGTGTTTTTAAGCCAATNANAAAATTTTTTNATGTTGTTTNNNNNNNNNNNNNNNNNNNNNN
CCACAAAATTCGGTTANTNTTTTAAAAAANTACAACAAANNNNNNNNNNNNNNNNNNNNNN

4381 NNN
NN

4441 NNN
NN

4501 NNN
NN

4561 NNN
NN

4621 NNN
NN

4681 NNN
NN

4741 NNN
NN

4801 NNN
NN

4861 NNN
NN

4921 NNN
NN

4981 NNN
NN

5041 NGCCANGRAGGGGGCCAGGTTCCAANTTCCCAACCKTTTTWGGARGGACNGCCCCAGGG
NCGGTNCYTCCCCCGGTCCAAGGTTNAAGGGTTGMAAAAWCCTYCCTGNCGGGGGTCCC

5101 GGGGATRAACAGANTNGGGGGKGGTWGGGTTNAGGGTGGAACNCCTTNGCGCCTGGAG
CCCCAYTTGTCTNANCCCCMCCAWCCCAANTMCCACCCTTGNGGAANCGSCGGACCTC

5161 AACGTGCAAAGAGGAAATGAAGGGCCTGKGTCAAGGAGCCCAAGTNGGCGGGGRAGTTTG
TTGCACGTTTCTCCTTTACTTCCCGGACMCAGTTCTTCGGGTTTCANCCGCCCCYTCAAAC

5221 CAGGGAGGCACTCCGGGGAGGTCCSGCGTGCCCGTCCAAGGGAGCAATGCGTCCTTCGGG
GTCCCTCCGTGAGGCCCCCTCAGGSCGCACGGGCAGGTTCCCTCGTTACGCAGGAAGCCC

5281 TTCGTCCCCAWGCCGCGTCTACGCGCCTYCCGTCTCCTCCCTTCACGTTCCGGCATTTCGTG
AAGCAGGGGTWCGGCGCAGATGCGCGGARGGCAGGAGGGGAAGTGCAAGGCCGTAAGCAC

5341 GTGCCCCGAGCCCGACGCCCCGCGTCCGGACCTGGAGGCAGCCCTGGGTCTCCGGATCAG
CACGGGCCCTCGGGCTGCGGGGCGCAGGCCTGGACCTCCGTGCGGACCCAGAGGCCTAGTC

5401 GCCAGCGGCCAAAGGGTCGCCGCACGCACCTGTTCCAGGGCCTCCACATCATGGCCCCCT
CGGTGCGCGGTTTCCAGCGGCGTGCGTGGAACAAGGGTCCCGGAGGTGTAGTACCGGGGA

FIG. 21
(CONTINUED)

5461 CCCTCGGGTTACCCACAGCCTAGGCCGGATTGACCTCTCTCCGCTGGGGCCCTCGCCT
GGGAGCCCAATGGGGTGTGCGATCCGGCCTAAGCTGGAGAGAGGCGACCCCGGGAGCGGA

Sp1

5521 GGCGTCCCTGCACCCTGGGAGCGCGAGCGGCGCGGGCGGGGAAGCGCGGCCCATACCC
CCGACGGGACGTGGGACCCTCGCGCTCGCCGCGCGCCCGCCCTTCGCGCCGGGTATGGG

5581 CCGGGTCCGCCCCGAAGCAGCTGCGCTGTGCGGGCCAGGCCGGGCTCCCAGTGGATTTCG
GGCCAGGCGGGCCTTCGTGACGCGACAGCCCCGGTCCGGCCCCGAGGGTCACCTAAGCG

Topo_II_cleavage_site

5641 GGGCACAGACGCCCAGGACCGCGCTTCCACGTGGCGGAAGGACTGGGGACCCGGGCACC
CCCGTGTCTGCGGGTCTTGGCGCGAAGGGTGCACCGCCTTCCTGACCCCTGGGCCCCGTGG

E2F

5701 CGTCCTGCCCCCTTCACCTTCCAGCTCCGCTTCTTCCGCGCGGACCCGGCCCCGTCCCGAA
GCAGGACGGGGAAGTGAAGGTGAGGCGAAGAAGGCGCGCTGGGCGGGGCAGGGCTT

E

5761 CCCTTCCCAGGTCCCGGCCCCAGCCCCCTCCGGGCCCTCCCAGCCCCCTCCCTTCTTTTC
GGGAAGGGTCCAGGGCCGGGTGCGGGAAGGCCCGGGAGGGTCGGGGAGGGGAAGGAAAAG

Sp1
=====

2F NFkB h

5821 CGCGGCCCCGCCCTCTCCTTCGCGGCGCGAGTTTCAGGCAGCGCTGCGTCTGCTGCGCA
GCGCCGGGGCGGGAGAGGAAGCGCCGCGCTCAAAGTCCGTCGCGACGCAGGACGACGCGT

5860 5875
ECO47III FSP1

TRT5 '
*****>

5881 CGTGGGAAGCCCTGGCCCCGCCACCCCGCGATGCCGCGCGCTCCCGCTGCCGAGCCG
GCACCCTTCGGGACCGGGGCCGGTGGGGGCGCTACGGCGCGCGAGGGGCGACGGCTCGGC

5941 TGCGTCCCTGCTGCGCAGCCACTACCGCGAGGTGCTGCCGCTGGCCACGTTGCTGCGGC
ACGCGAGGGACGACGCGTCCGTGATGGCGCTCCACGACGGCGACCGGTGCAAGCACGCCG

5953
FSP1

6001 GCCTGGGGCCCCAGGGCTGGCGGCTGGTGCAGCGCGGGGACCCGGCGGCTTTCCGCGCGC
CGGACCCCGGGGTCCCGACCGCCGACCACGTGCGCCCCCTGGGCCGCCGAAAGGCGCGCG

6061 TGGTGGCCCAGTGCTTGGTGTGCGTGCCCTGGGACGCACGGCCGCCCCCGCCGCCCCCT
ACCACGGGTACAGGACCACACGCACGGGACCTGCGTGCCGCGGGGGGCGGCGGGGGA

NFkB
=====

FIG. 21
(CONTINUED)

```
-----
*****
6121 CCTTCCGCCAGGTGGGCTCCCGGGGTCGGCTCCGGCTGGGGTTGAGGGCGCCGGGG
    GGAAGGCGGTCCACCCGAGGGGCCAGCCGACGCCGACCCCAACTCCCGCCGGCCCC

                                Topo_II_cleavage_s
                                ::::::::::::::
                                NFkB
                                =====

    Intron1
    *****>
6181 GGAACCAGCAGACATGCGGAGAGCAGCGCAGGCGACTCAGGGCGCTTCCCCCGCAGGTGTC
    CCTTGGTCGCTGTACGCCCTCTCGTCGCGTCCGCTGAGTCCCGCAAGGGGGCGTCCACAG

    ite
    :

6241 CTGCCTGAAGGAGCTGGTGGCCCGAGTGCTGCAGAGGCTGTGCGAGCGCGGCGCAAGAA
    GACGGACTTCCTCGACCACCGGGCTCACGACGTCTCCGACACGCTCGCGCCGCGCTTCTT

6301 CGTGCTGGCCTTCGGCTTCGCGCTGCTGGACGGGGCCCGGGGGCCCCCGAGGCCTT
    GCACGACCGGAAGCCGAAGCGCGACGACCTGCCCCGGGCGCCCCGGGGGGGCTCCGGAA

6361 CACCACCAGCGTGCGCAGCTACCTGCCAACACGGTGACCGACGCACTGCGGGGGAGCGG
    GTGGTGGTTCGCACGCGTCGATGGACGGGTGTGCCACTGGCTGCGTGACCCCCCTCGCC
        ^
        6372
        FSP1

6421 GCGGTGGGGGCTGCTGCTGCGCCGCGTGGGCGACGACGTGCTGGTTACCTGCTGGCAGC
    CCGCACCCCCGACGACGACGCGGCGCACCCGCTGCTGCACGACCAAGTGGACGACCGTGC

6481 CTGCGCGCTCTTTGTGCTGGTGGCTCCCAGCTGCGCCTACCAGGTGTGCGGGCCGCCGCT
    GACGCGCGAGAAACACGACCACCGAGGGTCGACGCGGATGGTCCACACGCCCGGCGGCGA

6541 GTACCAGCTCGGCGCTGCCACTCAGGCCCGGCCCCCGCCACACGCTAGTGGACCCCGAAG
    CATGGTCGAGCCGCGACGGTGAGTCCGGGCGGGGGCGGTGTGCGATCACCTGGGGCTTC

6601 GCGTCTGGGATGCGAACGGGCCTGGAACCATAGCGTCAGGGAGGCCGGGGTCCCCCTGGG
    CGCAGACCCCTACGCTTGCCCGGACCTTGGTATCGCAGTCCCTCCGGCCCCAGGGGGACCC

6661 CCTGCCAGCCCCGGGTGCGAGGAGGCGCGGGGGCAGTGCCAGCCGAAGTCTGCCGTTGCC
    GGACGGTCGGGGCCACGCTCCTCCGCGCCCCCGTACGGTCGGCTTCAGACGGCAACGG

6721 CAAGAGGCCCAGGCGTGGCGCTGCCCTGAGCCGGAGCGGACGCCCGTTGGGCAGGGGTC
    GTTCTCCGGGTCCGCACCGCGACGGGGACTCGGCCTCGCCTGCGGGCAACCCGTCCCCAG

6781 CTGGGCCCCACCGGGCAGGACGCGTGGACCGAGTGACCGTGGTTTCTGTGTGGTGTACCC
    GACCCGGGTGGGCGCGTCTGCGCACCTGGCTCACTGGCACCAAGACACACCACAGTGG

6841 TGCCAGACCCGCCGAAGAAGCCACCTCTTTGGAGGGTGCGCTCTCTGGCACGCGCCACTC
    ACGGTCTGGGCGGCTTCTTCGGTGGAGAACTCCACGCGAGAGACCGTGCGCGGTGAG

6901 CCACCCATCCGTGGGCCGCCAGCACCACGCGGGCCCCCATCCACATCGCGGCCACCACG
    GGTGGGTAGGCACCCGGCGGTGCTGGTGCGCCCGGGGGTAGGTGTAGCGCCGGTGGTGC
```

FIG. 21
(CONTINUED)

6961 TCCCTGGGACACGCCTTGTCCCCCGGTGTACGCCGAGACCAAGCACTTCCTCTACTCCTC
AGGGACCCTGTGCGGAACAGGGGGCCACATGCGGCTCTGGTTTCGTGAAGGAGATGAGGAG

7021 AGGCGACAAGGAGCAGCTGCGGCCCTCCTTCCTACTCAGCTCTCTGAGGCCCAGCCTGAC
TCCGCTGTTCCTCGTCGACGCCGGGAGGAAGGATGAGTCGAGAGACTCCGGGTTCGGACTG

7081 TGGCGCTCGGAGGCTCGTGGAGACCATCTTTCTGGGTTCAGGCCCTGGATGCCAGGGAC
ACCGCGAGCCTCCGAGCACCTCTGGTAGAAAGACCCAAGGTCCGGGACCTACGGTCCCTG

7141 TCCCCGCAGGTTGCCCCGCCTGCCCCAGCGCTACTGGCAAATGCGGGCCCCTGTTTCTGGA
AGGGGCGTCCAACGGGGCGGACGGGTTCGCGATGACCGTTTACGCCGGGGACAAAGACCT

7167
ECO47III

7201 GCTGCTTGGGAACACGCGCAGTGCCCCCTACGGGGTGCTCCTCAAGACGCACTGCCCCGCT
CGACGAACCCTTGGTGCGCGTCACGGGGATGCCCCACGAGGAGTTCTGCGTGACGGGGCA

7261 GCGAGCTGCGGTACCCCCAGCAGCCGGTGTCTGTGCCCGGGAGAAGCCCCAGGGCTCTGT
CGCTCGACGCCAGTGGGGTCGTGCGCCACAGACACGGGCCCTCTTCGGGGTCCCGAGACA

7321 GCGGGCCCCCGAGGAGGAGGACACAGACCCCCGTCGCCTGGTGCAGCTGCTCCGCCAGCA
CCGCCGGGGGCTCCTCCTCCTGTGTCTGGGGGACGCGGACCACGTTCGACGAGGCGGTCTGT

7381 CAGCAGCCCCCTGGCAGGTGTACGGCTTCGTGCGGGCCTGCCTGCGCCGGCTGGTGCCCCC
GTCGTGCGGGACCGTCCACATGCCGAAGCACGCCCGGACGGACGCGGCCGACCACGGGGG

7441 AGGCCTCTGGGGCTCCAGGCACAACGAACGCCGCTTCCTCAGGAACACCAAGAAGTTCAT
TCCGGAGACCCCCAGGTCCGTGTTGCTTGC GGCGAAGGAGTCCTTGTGGTTCTTCAAGTA

7501 CTCCCTGGGGAAGCATGCCAAGCTCTCGCTGCAGGAGCTGACGTGGAAGATGAGCGTGCG
GAGGGACCCCTTCGTACGGTTCGAGAGCGACGTCCTCGACTGCACCTTCTACTCGCACCG

7561 GGA CTGCGCTTGGCTGCGCAGGAGCCCAGGTGAGGAGGTGGTGGCCGTCGAGGGCCCAGG
CCTGACGCGAACCGACGCGTCCTCGGGTCCACTCCTCCACCACCGGCAGCTCCCGGGTCC

7575
FSP1

Intron2

7621 CCCCAGAGCTGAATGCAGTAGGGGCTCAGAAAAGGGGGCAGGCAGAGCCCTGGTCCTCCT
GGGGTCTCGACTTACGTCATCCCCGAGTCTTTTCCCCCGTCCGTCTCGGGACCAGGAGGA

7681 GTCTCCATCGTCACGTGGGCACACGTGGCTTTTCGCTCAGGACGTCGAGTGGACACGGTG
CAGAGGTAGCAGTGCAACCGTGTGCACCGAAAAGCGAGTCCTGCAGCTCACCTGTGCCAC

7741 ATCGAGGTCGACTCTAGAGGATCCCCGGGTACCGAGCTCGAATTCGTAATCATGGTCATA
TAGCTCCAGCTGAGATCTCCTAGGGGCCCATGGCTCGAGCTTAAGCATTAGTACCAGTAT

7747
SAL1

FIG. 21
(CONTINUED)

32/34

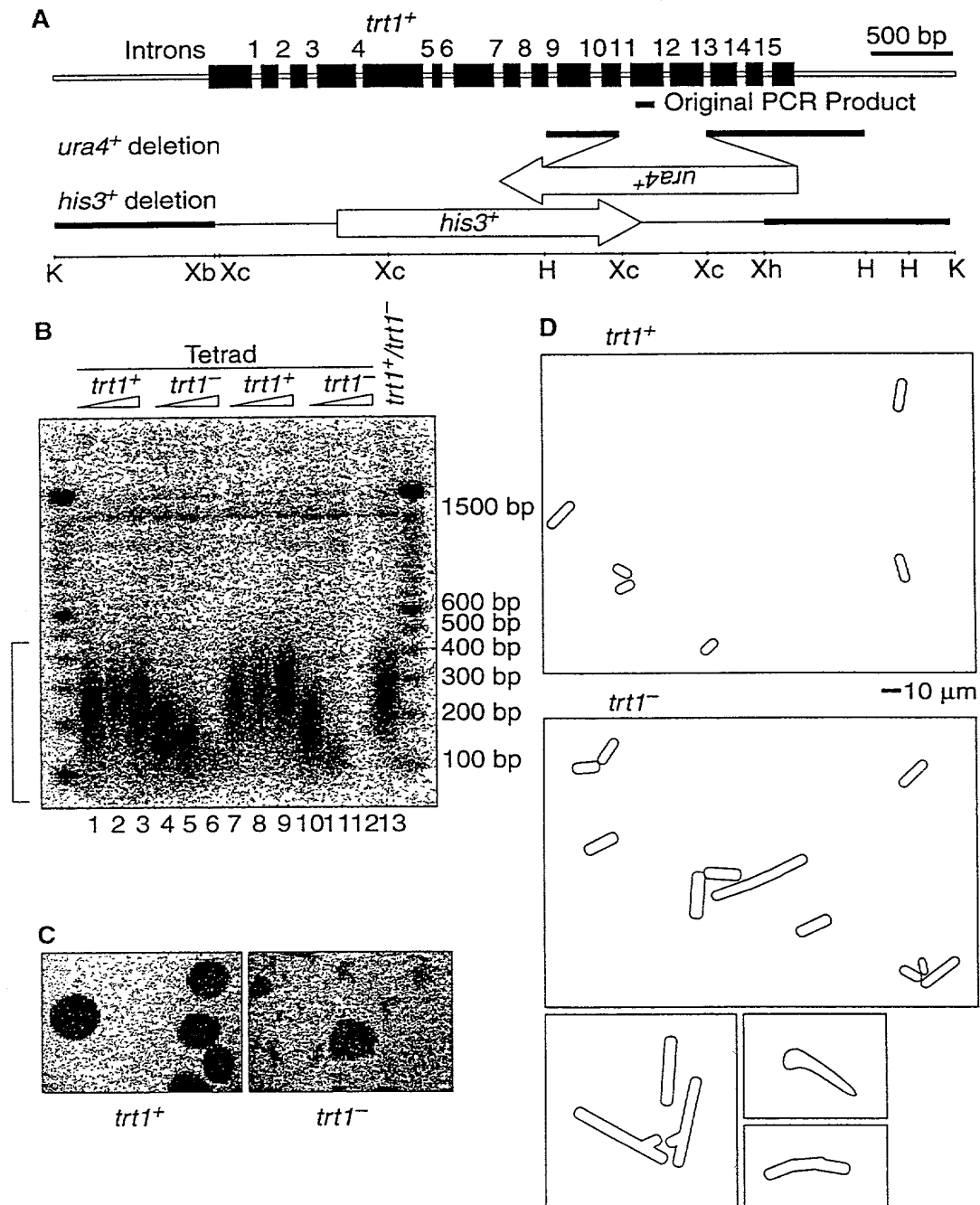


FIG. 22

33/34

gccaagttcctgcactggctgatgagtgtgtacgtcgctcgagctgctcaggtctttcttt
tatgtcacggagaccacgtttcaaaagaacaggctctttttctaccggaagagtgtctgg
agcaagttgcaaagcattggaatcacacagcacttgaagaggggtgcagctgcgggacgtg
tcggaagcagaggtcaggcagcatcggaagccaggcccgccctgctgacgtccagactc
cgcttcatccccaagcctgacgggctgcggccgattgtgaacatggactacgtcgtggga
gccagaacgttccgcagagaaaaagagggccgagcgtctcacctcgaggggtgaaggcactg
ttcagcgtgctcaactacgagcgggcgcg

FIG. 23

TCTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGCTCACCTGCAGGAG
ACCAGCCCCGCTGAGGGATGCCGTCGTCATCGAGCAGAGCTCCTCCCTGAATGAGGCC
AGCAGTGGCCTCTTCGACGTCCTTCCTACGCTTCATGTGCCACCACGCCGTGCGCATC
AGGGGCAAGTC

FIG. 24

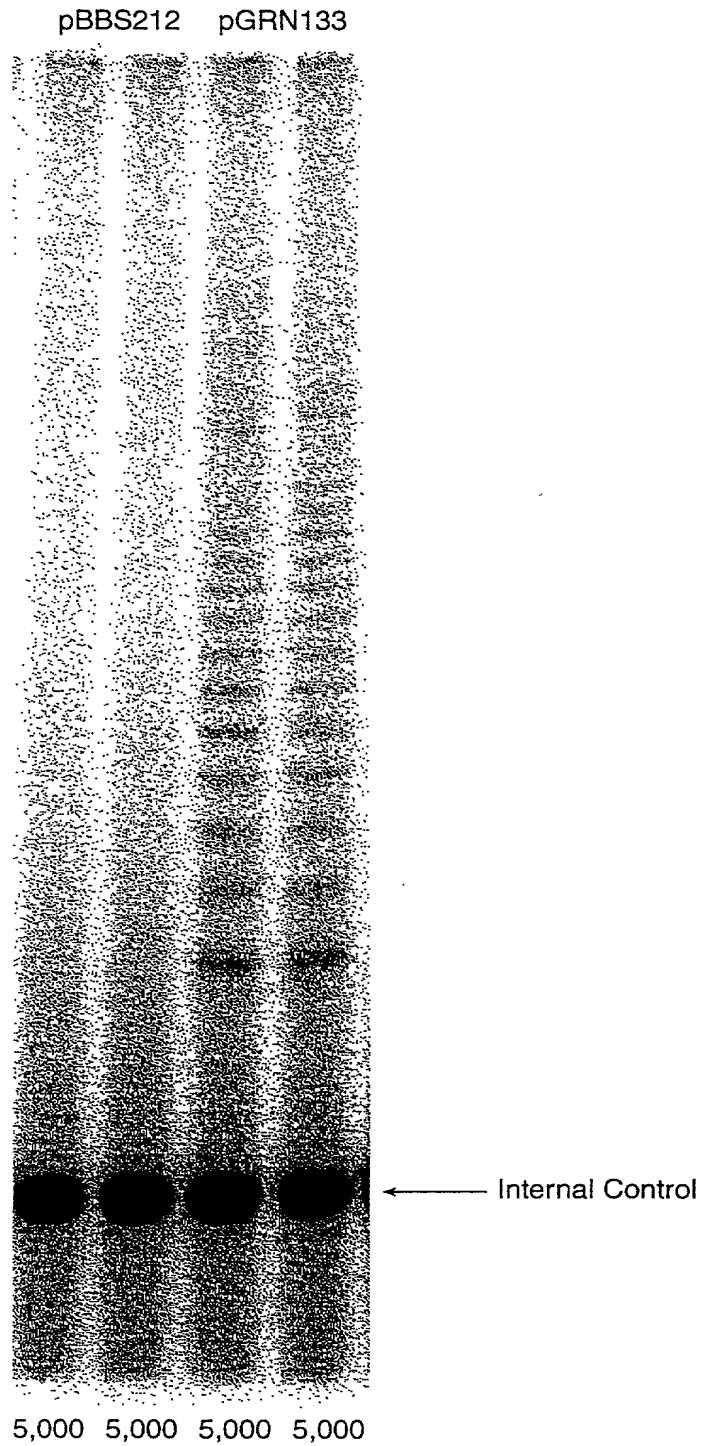


FIG. 25